行政院國家科學委員會專題研究計畫精簡報告

以HCCA在無線區域網路提供變動位元速率訊務之服務品質保證(1/2)
HCCA Based Quality of Service Guarantee for VBR Traffic in Wireless LANs
計畫編號：NSC 98-2221-E-009-062-MY2
執行期間：98/08/01 ~99/07/31

主持人：李程輝教授 國立交通大學電信工程研究所
參與人員：黃郁文、謝景融、黃迺倫、林建碩、机奕璉、黃煜傑 國立交通大學電信工程研究所

中文摘要

在無線區域網路標準(IEEE 802.11e)中提供了一種在中控型通道擷取(HCCA)模式下的排程器設計。在此設計中，頻寬分配根據訊務平均資料量，因而對於變動位元速率的訊務，可能出現封包的違反延遲限制。過去的文獻建議透過工作站佇列狀態回報的機制，使得系統能分配足夠的頻寬杜絕封包違反延遲限制的發生，然而，及時的佇列回報可能位系統帶來可觀的額外負擔從而降低頻寬使用效率，在本研究中，我們將提出一種基於多重輪詢技術的低負擔佇列狀態回報機制並將其結合最早時限優先的排程原則於無線區域網路中提供服務品質保證。電腦模擬實驗結果顯示，相較於先前的研究結果，本研究提出的方法不僅能夠達成服務品質的需求並且更有效的管理頻寬。

關鍵字：服務品質，無線區域網路、佇列狀態回報

***Abstract*—HCF Controlled Channel Access (HCCA), defined in the IEEE 802.11e document, provides a sample scheduler to allocate transmission opportunity (TXOP) to QoS-aware stations (QSTAs). Since the calculation of TXOP duration is based on mean data rate, it is efficient for constant bit rate (CBR) traffic. For variable bit rate (VBR) traffic, however, delay bound of some admitted traffic flows may be violated. Several previous works suggested that, through reporting queue status to schedulers, each QSTA can be allocated sufficient transmission opportunity (TXOP) for its buffered packets to meet delay bound requirement. However, timely queue status reports may cause considerable overheads and lead to inefficient bandwidth utilization. In this paper, we propose a low overhead queue status report scheme based on multi-polling technique and earliest deadline first policy (EDF) to provide QoS support. Simulation results show that our proposed scheme can not only meet the QoS requirements but also manage the bandwidth more efficiently than previous works.**

***Index Terms* – QoS, WLAN, Queue status report**

I. INTRODUCTION

Wireless Local Area Network (WLAN) technology [1] offers a simple and convenient solution to ubiquitous wireless access. As multimedia applications become more and more popular, the capability to provide QoS support over WLAN becomes an important research issue. IEEE 802.11e standard [2] introduces a new coordination function, namely, hybrid coordination function (HCF), which consists of two medium access mechanisms. One is contention-based Enhanced Distributed Channel Access (EDCA), and the other is contention-free HCF Controlled Channel Access (HCCA).

HCCA requires a centralized QoS-aware coordinator, called Hybrid Coordinator (HC), to manage the medium access. HC can poll normal QoS-aware stations (QSTAs) after sensing the medium idle for a PCF inter-frame space (PIFS) that is shorter than DCF inter-frame space (DIFS) adopted by QSTAs. Therefore, HC have higher priority than QSTAs to access the medium. After gaining the control of the transmission medium, HC will poll QSTAs according to its polling list. In order to be included into the polling list, each QSTA needs to make a separate QoS service reservation, which can be accomplished by sending Add Traffic Stream (ADDTS) frame to HC. In this frame, QSTAs can describe the detailed traffic characteristics and QoS requirements in the Traffic Specification (TSPEC) field. Based on these traffic characteristics and the QoS requirements, HC determines the common scheduled service interval (SI) and transmission opportunity (TXOP).

Upon receiving a poll, the polled QSTA either responds with QoS-Data if it has packets to send or QoS-Null frame otherwise. When the TXOP duration ends, HC gains the control of channel again and either sends QoS-Poll to next QSTA on its polling list or releases the medium if there is no QSTA to be polled.

The sample scheduler, provided in the IEEE 802.11e document, allocates TXOPs to QSTAs based on the mean data rate and nominal MSDU size. It performs well for constant bit rate (CBR) traffic. For variable bit rate (VBR) traffic, the requested delay bound may be violated, and packet loss may occur seriously.

To handle VBR traffic flows, prediction and optimization-based HCCA (PRO-HCCA) [3] request each QSTA to piggyback its queue status in the headers of data frames. Another scheme, two-step multi-polling (TS-MP) [4], initiates a separate period for queue status collection at the beginning of each SI. Based on the reported queue status and the requested delay bound, HC can wisely allocate TXOP to QSTAs to reduce the delay bound violation probability.

However, for real-time VBR traffic, timely queue status report may require considerable overheads, leading to inefficient bandwidth utilization. These may result from the need to increase SI or to report queue status separately. Due to the queue status obtained from piggyback process, when allocating TXOP to QSTAs, the PRO-HCCA scheduler can not obtain the amount of packets arrived in previous SI. To guarantee the delay bound requirements, the PRO-HCCA scheduler should set the SI no larger than half minimum of all requested delay bounds. On the other hand, since a time period is initiated to collect the queue status at the beginning of each SI, TS-MP can obtain the traffic amount of each flow arrived in previous SI and thus, the SI can be set larger than that chosen by the PRO-HCCA scheduler. Some unnecessary overheads, however, are caused if there are traffic flows with different delay bound in the system. The traffic flows with larger delay bounds may not have to separately report their queue status but only piggyback them as the way adopted by the PRO-HCCA scheduler.

Motivated by the above observations, in this paper, we propose an efficient queue status report scheme for TS-MP with combining the advantages of the piggyback process adopted by PRO-HCCA scheduler. Also, through obtaining the queue status and QoS requirements of traffic flows, we can efficiently schedule the transmission according to earliest deadline first (EDF) policy and effectively manage the queues based on the requested packet loss probability requirement. Simulation results show that our proposed scheme can not only meet the QoS requirements of each traffic flow but also, requires less overhead than the sample scheduler and PRO-HCCA. The sample scheduler, PRO-HCCA and TS-MP are related to our work and will be reviewed in Section III.

The rest of the paper is organized as follows. Section II defines the system model. The sample Scheduler, PRO-HCCA and TS-MP are reviewed in Section III. Section IV presents our proposed scheme. The simulation results are shown and discussed in Section V. Finally, we draw the conclusions in Section VI.

II. SYSTEM MODEL

We assume that transmission over the wireless medium is divided into SIs and the duration of each SI, denoted by *SI*, is a sub-multiple of the length of a beacon interval T_b. Moreover, a SI is further divided into a contention period and a contention-free period.

When the connection request of a new real-time flow arrives, a QSTA has to negotiate with the HC for admission. The QSTA needs to describe the traffic characteristics of the new flow in the TSPEC field of the ADDTS frame. As an example, the mandatory traffic characteristics defined in the consensus proposal include Mean Data Rate (ρ), Nominal MSDU Size (L) and Maximum Service Interval ($SI_{\max}$). The HC uses the specified traffic characteristics and the requested QoS requirements to calculate the TXOP allocation for the new flow and accepts it if the requested QoS can be guaranteed without violating the QoS requirements of existing connections.

In this paper, we assume that the QoS requirements are specified with delay bound and packet loss probability. The delay bound can be specified in the $SI_{\max}$ field, and the packet loss probability can be specified in some unused field of the ADDTS frame. The studied system consists of *N* traffic flows, called F_1, F_2,…, and F_N. Each flow requires a separate queue in its attached QSTA. Moreover, the queue, delay bound and packet loss probability of F_i are denoted, respectively, by $Queue_i$, D_i and P_i.

III. RELATED WORK

In the section, we briefly review the sample scheduler, PRO-HCCA and TS-MP scheme.

A. The Sample Scheduler

Let ρ_i, L_i, and $SI_{\max,i}$ denote, respectively, the mean data Rate, nominal MSDU Size, and maximum Service Interval of F_i. Assume that a new traffic stream, i.e., F_{N+1}, is requesting for admission. In the sample scheduler, the HC determines a possible new *SI* according to $SI = \min\{SI, SI_{\max,N+1}\}$.

The sample scheduler calculates the TXOP duration for F_i by the following steps. Firstly, the sample scheduler decides the average number of packets N_i that arrives at the mean data rate during one *SI* for F_i:

$$N_i = \left\lceil \frac{\rho_i \cdot SI}{L_i} \right\rceil. \tag{1}$$

Secondly, the TXOP duration is obtained for F_i as follows:

$$TD_i = \max\left(N_i \cdot \left(\frac{L_i}{R_i} + O \right), \frac{M_i}{R_i} + O \right) \tag{2}$$

where R_i is minimum physical transmission rate of the attached QSTA of F_i, L_i is nominal MSDU Size, M_i is maximum MSDU size, and O denotes the per-packet overhead (which includes the transmission time for ACK frame, inter-frame space, MAC header, CRC field and PHY PLCP Preamble and Header).

Again, the sample scheduler is only suitable for CBR traffic. For VBR traffic, it may cause serious delay bound violation because of the fluctuation of data rate and packet size.

B. PRO-HCCA

To handle VBR traffic flows, the PRO-HCCA scheduler requests each QSTA to piggyback its queue status in the header of frames transmitted in the allocated TXOP. In addition, to treat packets from different traffic flows with their urgencies for services, for F_i, a partition list, PL_i, is maintained with entry *j* (i.e., $PL_{i,j}$) to record the amount of packets backlogged for time period between $(j-1)\cdot SI$ and $j \cdot SI$, $1 \le i \le N$, $1 \le j \le \lceil D_i/SI \rceil$. When allocating TXOP to QSTAs, the PRO-HCCA scheduler can not obtain $PL_{i,1}$, $1 \le i \le N$, by piggyback process on time. It is predicted by adopting wavelet least mean square (WLMS) predictor [5], [6]. Also, the PRO-HCCA scheduler defines a utility function, $U_{ij} = 1/(\lceil D_i/SI \rceil - j + 1)$, which represents the utility received for transmitting packets belonging to $PL_{i,j}$ for a single time unit. Let R_i, T_{avail} and t_{ij} represent, respectively, the adopted physical transmission rate of the attached QSTA of F_i, the available time reserved for the

PRO-HCCA scheduler and the allocated transmission time for $PL_{i,j}$, $1 \le i \le N$, $1 \le j \le \lceil D_i/SI \rceil$. Through maximizing

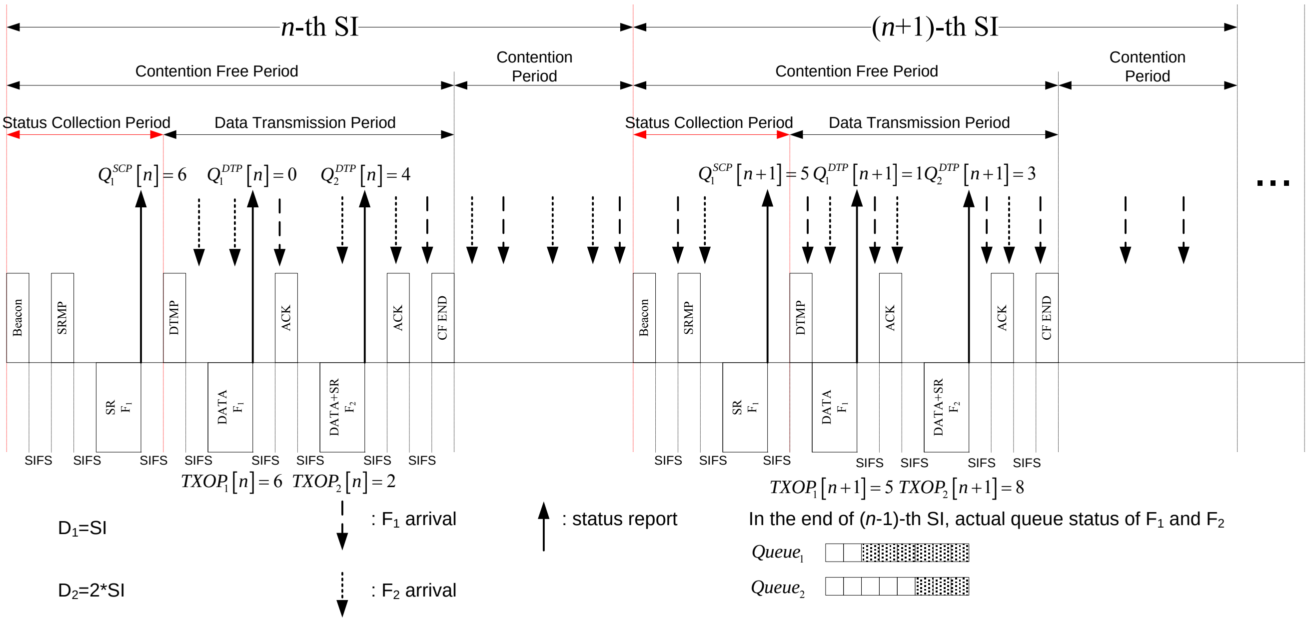

Fig.1 An example for illustrating our proposed queue status report scheme.

$$\sum_{i=1}^{N}\sum_{j=1}^{\lceil D_i/SI \rceil} U_{ij}\cdot t_{ij} \tag{3}$$

subject to $\sum_{i=1}^{N}\sum_{j=1}^{\lceil D_i/SI \rceil} t_{ij} \le T_{avail}$ and $0 \le t_{ij} \le \left(PL_{i,j}/R_i\right)+O$, t_{ij} can be calculated [7]. Furthermore, the allocated TXOP for F_i, $TXOP_i$, is determined as follows

$$TXOP_i = \sum_{j=1}^{\lceil D_i/SI \rceil} t_{ij}\ . \tag{4}$$

Since the predicted value, $PL_{i,1}$, and the actual value may be mismatched, the PRO-HCCA scheduler can provide delay guarantee at most up to $2SI$, resulting in increase of polling overheads.

A. TS-MP

Another design, called two-step multi-polling (TS-MP) scheme [4], can also handle VBR traffic flow through queue status report. It divides the contention-free period into two parts. One is status collection period (SCP), and the other one is data transmission period (DTP). In the beginning of each SI, access point (AP) will initiate SCP by broadcasting status request multi-poll (SRMP) frame, which describes which QSTAs should report their queue status. Upon receiving SRMP, each polled QSTA describes its buffer status in status response (SR) frame and send it back in turn. After receiving the last SR frame, AP will start DTP with broadcasting data transmission multi-poll (DTMP) frame, which schedules the transmission order and the allocated TXOP for each polled QSTA. Similarly, upon receiving this frame, QSTAs will transmit packets based on the specified schedule.

IV. OUR PROPOSED SCHEME

In this section, we present in detail the low overhead queue status report scheme for TS-MP and the scheduling algorithm derived from EDF policy.

A. Low Overhead Queue Status Report Scheme

Firstly, we decide SI as

$$SI = \min_{1\le i\le N} D_i \tag{5}$$

To avoid delay bound violation and reduce the unnecessary overheads, in SCP, we only collect queue status separately from traffic flows with their delay bound satisfying $\lfloor D_i/SI \rfloor = 1$, $1 \le i \le N$. The collected queue status of F_i in the n-th SCP is represented by $Q_i^{SCP}[n]$. In DTP, each traffic flow F_i, $1 \le i \le N$, will transmit its packets with the buffer status piggybacked. Let $Q_i^{DTP}[n]$ and $TXOP_i[n]$ denote, respectively, the reported queue status and allocated TXOP of F_i in the DTP of the n-th SI. Note that $Q_i^{SCP}[n]$, $Q_i^{DTP}[n]$ and $TXOP_i[n]$ are all represented in time unit. It may occur that

$$Q_i^{DTP}[n] + TXOP_i[n] > Q_i^{SCP}[n], \tag{6}$$

meaning that some packets of F_i arrive during the time period between queue status reports in SCP and DTP. To reduce the delay bound violation probability, the partition list $PL_{i,1}$ for F_i is calculated as

$$PL_{i,1} = \begin{cases} \max\left(Q_i^{SCP}[n], E_i[n]\right) & \text{if } \lfloor D_i/SI \rfloor = 1 \\ E_i[n] & \text{if } \lfloor D_i/SI \rfloor > 1 \end{cases} \quad (7)$$

where $E_i[n]$ is the estimated traffic amount in time unit for packets of F_i, which arrived during time interval between the end of its previous TXOP and the start of its current one. Moreover, $E_i[n]$ is the output of *M*-order WLMS predictor with $(x_i[n-1], x_i[n-2], ..., x_i[n-M])$ as input, where

$$x_i[n-m] = Q_i^{DTP}[n-m] + TXOP_i[n-m] - Q_i^{DTP}[n-m-1]. \quad (8)$$

The rest of partition lists and the content of WLMS predictor are the same as those in original PRO-HCCA. Fig.1 shows an example of our proposed queue status report scheme. In this example, we assume that there are two traffic flows in the system. The requested delay bound of one traffic flow is twice of that of the other. Their queues contain six and three packets at the end of (*n*-1)-th SI, respectively.

B. EDF-based Scheduling Algorithm

Assume that $PL_{i,j}$, $1 \le i \le N$, $1 \le j \le \lfloor D_i/SI \rfloor$, is calculated. For convenience, we let $PL_{i,j} = 0$ for $j \le 0$ and $j > \lfloor D_i/SI \rfloor$, $1 \le i \le N$. Firstly, we determine if $\sum_{i=1}^{N} \sum_{j=1}^{\lfloor D_i/SI \rfloor} PL_{i,j} < T_{avail}$. If the condition holds, each traffic flow can obtain its TXOP as follows

$$TXOP_i = \sum_{j=1}^{\lfloor D_i/SI \rfloor} PL_{i,j} + O, \ 1 \le i \le N. \quad (9)$$

Note that *O* denotes the necessary overheads. Otherwise, decide the minimum *J* such that

$$\sum_{i=1}^{N} \sum_{j=0}^{J} PL_{i,(\lfloor D_i/SI \rfloor - j)} \ge T_{avail} \quad (10)$$

Define

$$Loss = \sum_{i=1}^{N} \sum_{j=0}^{J} PL_{i,(\lfloor D_i/SI \rfloor - j)} - T_{avail} \quad (11)$$

If $J = 0$, some packets belong to $PL_{i,\lfloor D_i/SI \rfloor}$ may violate their delay bound, and all packets of $PL_{i,(\lfloor D_i/SI \rfloor - j)}$, $1 \le j \le \lfloor D_i/SI \rfloor$, $1 \le i \le N$, can not be transmitted in this SI. If $J > 0$, some packets of traffic flows remain in the buffer for at least one more SI. Therefore, TXOP allocation for F_i can be based on the requested mean data rate ρ_i and packet loss probability P_i such that

$$TXOP_i = \begin{cases} \sum_{j=0}^{J} PL_{i,\lfloor D_i/SI \rfloor - j} - \dfrac{Loss \cdot P_i \cdot \rho_i}{\sum_{k \in \Lambda} P_k \cdot \rho_k} + O \ \ if \ \ i \in \Lambda \\ \sum_{j=0}^{J-1} PL_{i,\lfloor D_i/SI \rfloor - j} + O \qquad otherwise \end{cases}, \quad (12)$$

where Λ contains the traffic flows such that $PL_{i,\lfloor D_i/SI \rfloor - J} > 0$.

V. SIMULATION RESULTS

The PHY and MAC parameters and all related information used in simulations are shown in Table I and II. Note that the sizes of QoS-ACK and QoS-Poll in the table only include the sizes of MAC header and CRC overhead. We assume that the minimum physical rate is 2Mbps and t_{PLCP} is reduced to 96 μs.

The traffic is delivered from QSTAs to AP and the contention-free period occupies the whole SI. The studied system consists of two QSTAs. (i.e., QSTA 1 and QSTA 2) Each QSTA is attached with one real-time VBR flow, which are generated by video trace files [8], [9]. We consider two types of traffic flows i.e., Type I and Type II. The detailed information, including TSPEC parameters and QoS requirements of each type of traffic flow are summarized in Table III.

The sample scheduler, PRO-HCCA and our proposed scheme are investigated in the simulation. Besides, we consider TS-MP scheme with another scheduler which requests all QSTAs to report their queue statuses in SCP and allocates TXOP to them only based on the reported information. For simplicity, we denote this scheme as TS-MP in the following discussion. In PRO-HCCA, we consider two scenarios, which set SI to be 20 ms and 40ms, respectively. For the other three schemes, SI is assumed to be 40 ms. To investigate the overhead efficiency, we define overhead efficiency ratio as total required overheads in time unit over total data transmission time.

TABLE I
RELATED PARAMETERS USED IN SIMULATIONS.

SIFS	10 μs
MAC Header size	32 bytes
CRC size	4 bytes
QoS-AcK frame size	16 bytes
QoS CF-Poll frame size	36 bytes
PLCP Header Length	4 bytes
PLCP Preamble length	20 bytes
PHY rate(*R*)	11 Mbps
Minimum PHY rate ($R_{\min}$)	2 Mbps

TABLE II
TRANSMISSION TIME FOR DIFFERENT HEADER AND PER-PACKET OVERHEAD

PLCP Preamble and Header (t_{PLCP})	96 μs
Data MAC Header (t_{HDR})	23.2727 μs
Data CRC (t_{CRC})	2.90909 μs
ACK frame (t_{ACK})	107.63636 μs
QoS-CFPoll (t_{POLL})	122.1818 μs
Per-packet overhead (*O*)	249.81818 μs

TABLE III
TRAFFIC PARAMETERS FOR TYPE I AND II TRAFFIC FLOWS

Traffic Parameter	Type I	Type II
	QSTA 1	QSTA 2
	Non-interactive	Interactive

	video	video
	Jurassic Park I	Office Cam
Packet Loss Rate Requirement (P_L)	0.001	0.01
Maximum Service Interval ($SI_{\max}$)	40 (ms)	80 (ms)
Mean Data Rate (ρ)	268k(bps)	91k(bps)
Nominal MSDU size (L)	1339 (bytes)	452(bytes)

Fig. 2 and Fig. 3 show, respectively, the cumulative distribution function experienced by packets of QSTA 1 and QSTA 2. It is easily found that if SI is set to be 40 ms for all schemes, our proposed scheme can achieve delay bound violation probability less than 5%, while in PRO-HCCA and the sample scheduler, there are almost up to 70% packets which violate their delay bound. Although TS-MP and PRO-HCCA with SI equal to 20 ms can have comparable or better performance, their required overhead is much more than those caused in our proposed scheme. This is demonstrated in Fig. 4. In Fig. 4, we assume that there exists one Type I traffic flow in the system. The overhead efficiency ratio is investigated for various numbers of Type II traffic flows. As shown in Fig. 4, when the number of Type II traffic flow increases, the overhead efficiency ratio increases because of lower nominal MSDU size of Type II traffic flow. In addition, our proposed scheme requires less overheads than those caused by the other three schemes. The reasons are as follows. 1) Compared with PRO-HCCA with SI equal to 20ms, to meet the QoS requirement, our proposed scheme sets larger SI, meaning that the polling overhead can be reduced. 2) Compared with TS-MP, in our proposed scheme, QSTAs attached with Type II traffic flows do not have to separately report queue status in SCP but only piggyback them in DTP. 3) Compared with sample scheduler, our proposed scheme can benefit from the multi-polling effect such that the overhead can be reduced more as the number of Type II traffic flows increases. Note that the required overheads of PRO-HCCA with SI equal to 40 ms are identical to that of the sample scheduler.

VI. CONCLUSIONS

In this paper, we have proposed a low overhead queue status report scheme and a EDF-based scheduling algorithm. Performance comparisons of our proposed and other schemes are conducted through computer simulations. Results show that, with QoS requirements supported, our proposed scheme causes less overheads in contention-free period such that more available bandwidth can be left for transmission in contention period. Our proposed scheme and previous works focus on real-time traffic whose required delay bound is larger than their inter-arrival time. It is an interesting further research topic to develop an efficient queue status report scheme for real-time traffic with average inter-arrival time much larger than the requested delay bound.

REFERENCES

[1] IEEE 802.11 WG: IEEE Standard 802.11-1999, Part 11: Wireless LAN MAC and Physical Layer Specifications. Reference number ISO/IEC 8802-11: 1999(E), 1999.

[2] IEEE Std. 802.11e-2005, Part 11: Wireless LAN medium access control and physical layer specifications Amendment 8: medium access control (MAC) quality of service enhancements, Nov. 2005.

[3] M. M. Rashid, E. Hossain, V. K. Bharggava, " Controlled channel access scheduling for guaranteed QoS in IEEE 802.11e-based WLANs," IEEE Trans. Wireless Commun. vol. 7, no. 4, pp.1287-1297, Apr. 2008.

[4] B. S. Kim, S. Kim, Y. Fang, and T.F. Wong, "Two-Step Multipolling MAC Protocol for Wireless LANs," IEEE J. Selected Areas in Comm., vol. 23, no. 6, pp. 1276-1286, June 2005.

[5] B. Farhang-Boroujeny, Adaptive Filters: Theory and Applications. New York: Wiley, 1998.

[6] J. C. Lee and C. K. Un, "Performance of transform-domain LMS adaptive digital filters," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-34, pp. 499-510, 1986.

[7] T. Cormen, C. Leisorson, R. Rivest, and C. Stein, Introduction to Algorithms, Second Edition. Cambridge, MA: MIT Press, 2001.

[8] MPEG-4 and H.263 video traces for network performance evaluation, http://www.tkn.tu-berlin.de/research/trace/trace.html, Oct. 2006.

[9] F. Kozamernik, "Media streaming over the Internet – an overview of delivery technologies," EBU TECHNICAL REVIEW, Oct. 2002.

[10] H. Fattah and C. Leung, "An overview of scheduling algorithm in wireless multimedia networks," IEEE Wireless Commun., vol. 9, no. 5, pp. 76-83, Oct. 2002.

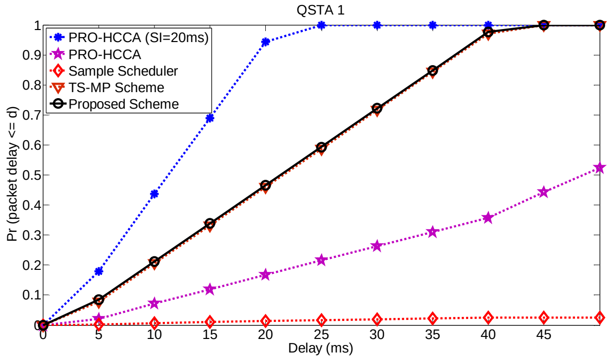

Fig.2 Cumulative distribution function of delay (ms) experienced by packets of QSTA 1

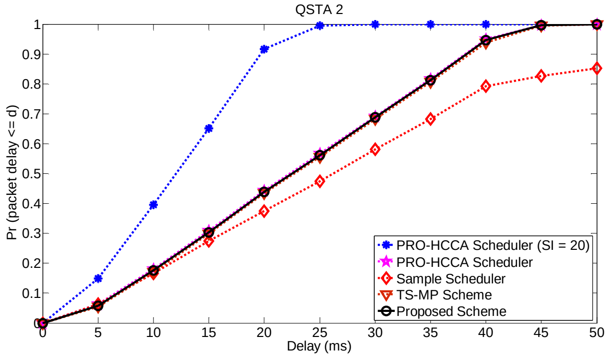

Fig.3 Cumulative distribution function of delay (ms) experienced by packets of QSTA 2

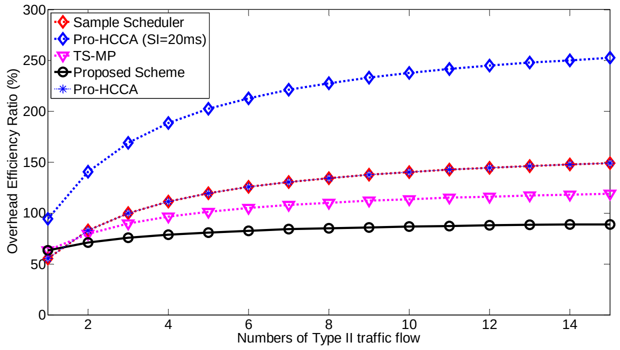

Fig. 4 Overhead efficiency ratio for different number of Type II traffic flow