行政院國家科學委員會補助專題研究計畫成果報告

※※※※※※※※※※※※※※※※※※※※※※※※※
※　　　　　　　　　　　　　　　　　　　　　　　　※
※　3G 無線接取網路之無線資源管理技術子計劃四　※
※　　　　無線封包網路之資源管理技術(I)　　　　※
※　　　　　　　　　　　　　　　　　　　　　　　　※
※※※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：□個別型計畫　　▒整合型計畫

計畫編號：NSC　NSC91-2219-E009-016

執行期間：　91 年　8　月　1　日至　92　年　7　月　31 日

計畫主持人：王蒞君副教授

本成果報告包括以下應繳交之附件：

□赴國外出差或研習心得報告一份

□赴大陸地區出差或研習心得報告一份

▒出席國際學術會議心得報告及發表之論文各一份

□國際合作研究計畫國外研究報告書一份

執行單位：國立交通大學電信工程系

中　華　民　國　92　年　7　月　31　日

成果報告(一)

寬頻分碼多重存取系統中高速向下連結封包存取的[1] 停滯防制機制之性能比較

Performance Comparison of Stall Avoidance Mechanisms for High Speed Downlink Packet Access in the WCDMA System

計畫編號：NSC 91-2219-E009-016
執行期限：91 年 8 月 1 日至 92 年 7 月 31 日

主持人：王蒞君副教授　國立交通大學電信工程系
計畫參與人員：張志文 張漢光 謝青洋　國立交通大學電信工程系

1 中文摘要

在這個計劃中，我們深入研究寬頻分碼多工存取（W-CDMA**）系統中高速向下聯結封包存取（**HSDPA: high speed downlink packet access**）的停滯防止**(stall avoidance)**機制。停滯防止的目的在減少傳輸延遲並且保持媒體存取控制（**MAC：Medium Access Control**）層的資料能夠依序的傳送到上層。我們藉由分析與模擬比較了三種停滯防止的機制，包含了以計時器（**timer-based**）、視窗（**window-based**）以及指示器（**Indicator-based**）為基準的停滯防止機制。為了方便比較，我們重新定義了一個新的成果指標稱為間斷處理時間（**gap processing time**）。我們的結果呈現出在高斯通道的環境中，當訊雜比（** E_b/N_0 **）為** 7 dB **時，計時器基準、視窗基準以及指示器基準的間斷處理時間分別為** 8.49、36.38 **以及** 175.5 **個傳輸時間間隔（**TTI: transmission time intervals**）。在瑞雷衰減通道環境中，我們發現指示器基準以及視窗基準的間斷處理時間分別都比高斯通道環境中的要高，而計時器基準的間斷處理時間則沒有明顯的差別。**

關鍵詞：**停滯防止，計時器基準，視窗基準，指示器基準。**

Abstract

In this project, we investigate stall avoidance mechanisms for the W-CDMA system with high speed downlink packet access (HSDPA). The stall avoidance mechanisms are aimed to reduce transmission delays and keep in-sequence delivery of MAC layer data to the upper layer. By analysis and simulation, we compare three stall avoidance mechanisms, including the timer-based, the window-based, and the indicator-based methods. For comparison, we introduce a new performance metric called the gap processing time. Our results show that for E_b/N_0 equal to 7 dB in the AWGN channel, the average gap processing time for the indicator-based method is 8.49 transmission time intervals (TTIs) and those for the window-based method and the timer-based method are 36.38 TTIs and 175.5 TTIs, respectively. In a Rayleigh fading channel, we find that the gap processing time of the window-based and the indicator-based methods are higher than in the AWGN channel, while the timer-based method has the same gap processing time in both the AWGN channel and the Rayleigh fading channel.

Keywords: stall avoidance, timer-based, window-based, indicator-based.

[1]**本文已在**IEEE PIMRC 2003 **刊登，詳細內容如附件。**

Performance Comparison of Stall Avoidance Mechanisms for High Speed Downlink Packet Access in the WCDMA System

Li-Chun Wang*, Chih-Wen Chang*, Han-Kuang Chang*, Chin-Yang Hsieh*, and Sam Jiang†
Department of Communication Engineering
*National Chiao Tung University, Taiwan
†AsusTek Computer Inc.
Email : lichun@cc.nctu.edu.tw

***Abstract*—In this paper, we investigate stall avoidance mechanisms for the W-CDMA system with high speed downlink packet access (HSDPA). The stall avoidance mechanisms are aimed to reduce transmission delays and keep in-sequence delivery of MAC layer data to the upper layer. By analysis and simulation, we compare three stall avoidance mechanisms, including the timer-based, the window-based, and the indicator-based methods. For comparison, we introduce a new performance metric called the gap processing time. Our results show that for $\frac{E_b}{N_0}$ equal to 7 dB in the AWGN channel, the average gap processing time for the indicator-based method is 18.55 transmission time intervals (TTIs) and those for the window-based method and the timer-based method are 42.56 TTIs and 170.8 TTIs, respectively. In a Rayleigh fading channel, we find that the gap processing time of the window-based and the indicator-based methods are higher than in the AWGN channel, while the timer-based method has the same gap processing time in both the AWGN channel and the Rayleigh fading channel.**

I. INTRODUCTION

HIGH speed downlink packet access (HSDPA) is becoming an important feature for the next generation wireless data networks. The third generation partnership project (3GPP) is also specifying the HSDPA standard for the wideband code division multiple access (WCDMA) system [1]. The goal of the WCDMA system with HSDPA is to achieve data rate up to 10 Mbits/sec in the mobile cellular environment. The key enabling technologies for the WCDMA system with HSDPA consist of fast link adaptation, fast scheduling, adaptive modulation and coding, and fast retransmission strategy in the MAC layer. As for efficient retransmission strategy in the MAC layer, the multi-channel stop-and-wait (SAW) hybrid automatic retransmission request (H-ARQ) is suggested to be implemented for HSDPA in the WCDMA system.

This work was supported jointly by the National Science Council R.O.C under the contract 90-2213-E-009-068, and 89-E-FA06-2-4.

However, the multi-channel SAW H-ARQ retransmission mechanism may encounter an undesirable stall situation in a wireless channel. The stall issue is defined as the situation when the receiver is waiting for the retransmission of a damaged packet data, while the transmitter believes that the packet has been transmitted successfully and never retransmit that packet. The stall issue causes long downlink transmission delays, which usually occurs in a wireless channel when a control signal is corrupted during the process of requesting packet retransmissions. For example, when the receiver detects a corrupted packet, it will send the negative acknowledgement (NACK) control signal to the transmitter to request retransmission. However, if the NACK control signal is changed to an acknowledgement (ACK) control signal, the transmitter may mistakenly believe that the packet was already successfully received. Then the transmitter removes the copy of this packet from its buffer and never send this packet again in the future. It has been reported that the probability of the NACK signal becoming the ACK signal may as high as 10^{-2} in a wireless channel [2]. Thus, it is critical to resolve the stall issue in order to reduce the downlink transmission delays and keep in-sequence delivery in wireless data networks.

In the literature, there are two research directions to resolve the stall issue in wireless data networks. The first research direction is to improve the reliability of the ACK and NACK control packets, e.g., by increasing the transmission power of control signals [2]. The second research direction to solve the stall issue is to design the stall avoidance mechanism in the MAC layer to detect the stall situation in the receiver [3], [4], [5], [6]. When a stall situation is detected, the receiver will send all the received packets within the packet sequences to the upper layers. Then the upper layer of the receiver will be responsible for requesting retransmission of the lost packets. In [3], the timer-based stall avoidance mechanism was introduced, which used a timer to control the operation of forwarding data packets to the upper layer. If the gap among the sequences of the received packets lasts over a predetermined expiration period, the timer-based mechanism will start forwarding the received packets to the upper layer. In [4], the window-based method proposed to use a sliding window to detect the stall situation before the timer expiration. In [5], [6], the indicator-based method introduced a concept of new data indicator (NDI) and the idea of mornitoring the activity of each H-ARQ process to activate the forwarding process of forwarding the received data to the upper layer before timer expiration and window overload.

The goal of this paper is to compare the performance of these three stall avoidance mechanisms by analysis and simulation in terms of the gap processing time, the new introduced performance metric. The gap processing time is defined as the period when a gap occurs in the sequences of the received packets until data packets in the reordering queue are delivered to the upper layer. Our resutls show that in an AWGN channel, the indicator-based mechanism performs the best. The window-based method performs worse than the indicator-based approach, but still much better than the timer-based method. In a Rayleigh fading channel, we find that the gap processing time of the window-based and the indicator-based methods are longer than that in AWGN channel, while the timer-based method has the same gap processing time in both the AWGN and Rayleigh fading channels.

The rest of this paper is organized as follows. In Section II, we introduce the multi-channel SAW H-ARQ mechanism and the stall problem in the MAC layer. Section III compare three stall avoidance mechanisms by analysis. Section IV shows the performance results through a cross-layer simulation under both the AWGN and the Rayleigh fading channels. Section V gives our concluding remarks.

II. THE STALL ISSUE IN THE MULTI-CHANNEL SAW H-ARQ

A. *Multi-Channel SAW H-ARQ*

The multi-channel SAW ARQ is used in the WCDMA system with HSDPA. The basic idea of the multi-channel SAW is to implement the concep of keeping the pipe full. Take a dual-channel SAW ARQ as an example, [7]. Figure 1 shows a dual-channel SAW ARQ consisting of two transmitters, i.e. the even transmitter and the odd transmitter. As shown in Figure 2, the even transmitter and the odd transmitter send the data alternatively. Specifically, after the even transmitter sends out data packet 0, it will wait the acknowledgement from the receiver. During this waiting period, the odd transmitter starts sending packet 1. Thus, the multi-channel SAW ARQ can achieve high throughput because it fully utilizes the channel capacity.

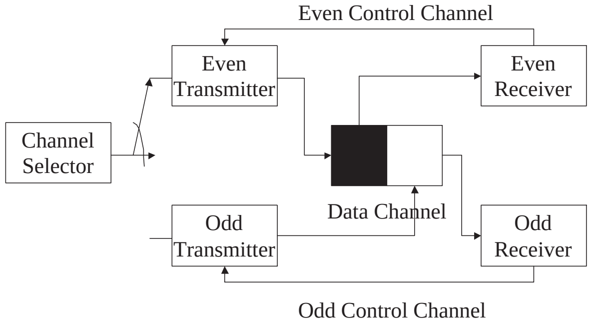

Fig. 1. The procedure of a dual-channel SAW H-ARQ scheme.

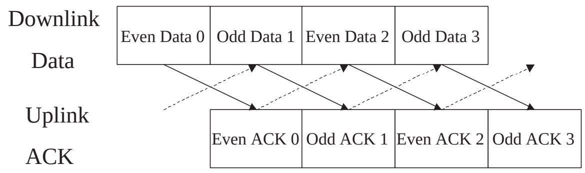

Fig. 2. The timing sequence of a dual-channel SAW H-ARQ scheme.

B. *Example of Stall Problem*

Although the multi-channel SAW H-ARQ can achieve high throughput. One of key issues to apply the multi-channel SAW H-ARQ in the wireless channel is the stall issue. The stall problem is defined as the case when the receiver is waiting a missing packet that will no longer be

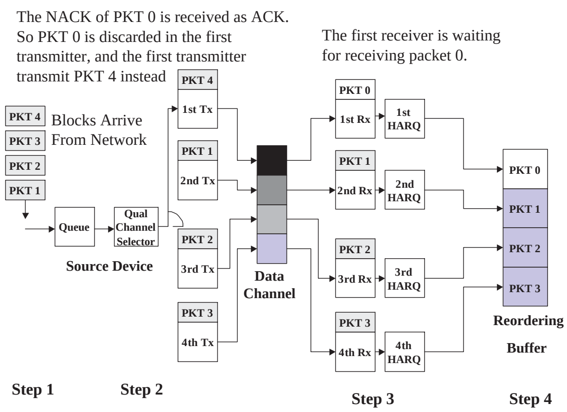

Fig. 3. A example of the stall problem in a qual-channel SAW H-ARQ scheme, where packet 0 is missing and packets 1,2, and 3 wait in the reordering buffer.

retransmitted by the transmitter.

Obviously, the stall problem in the MAC layer will cause long trnasmission delays to deliver data to the upper layer. The stall issue becomes very critical when the probability of NACK to ACK error is high. Figure 3 shows such an example in a qual-channel SAW H-ARQ. In the figure, packet 0 is missing. Thus, the first receiver sends NACK to the first transmitter through a feedback control channel. Assume an NACK to ACK error occurs during the transmission of NACK. Then the first transmitter thinks that packet 0 has already successfully reached the receiver, thereby removing the copy of packet 0 in the transmitter buffer and starting to send the next packet. In this stall situation, the first receiver will wait infinitely since packet 0 will no longer be sent again.

III. The Stall Avoidance Mechanism

Because control signals may be possibly corrupted during transmissions in a wireless channel, stall avoidance mechanisms are required to avoid the receiver waiting infinitely for data that will never be retransmitted. In this section, we discuss three stall avoidance mechanisms, including the timer-based, the window-based, and the indicator based schemes. Here we introduce a new performance metric-the gap processing time. The gap processing time is defined as the period starting when a gap appears in the sequences of the received packets until all the received data in the reordering queue are flushed to the upper layer even with some gaps.

A. Timer-Based Scheme

In [3], a timer-based stall avoidance mechanism was introduced. Based on the timer based approach, once a gap appears in the reordering buffer of the receiver, the timer is started to count how long the receiver has already waited for the retransmission of the missing packet. When the time expires, the reordering buffer stop waiting for the missing packet and deliver all the received data to the upper layers. Figure 4 shows an example of the operations in the timer based avoidance mechanism.

In Proposition 1, we analyze the average gap processing time for the timer-based stall avoidance method.

Proposition 1 *The average gap prossing time of the timer-based stall avoidance mechanism is $\frac{3}{2}T_{exp}$, where T_{exp} is the timer expiration time.*

Proof: Once a gap appears, the timer starts immediately. If there is only one gap in the reordering buffer, the gap processing time will be T_{exp} in a unit of the time transmission intervals (TTIs). However, if there are two or more gaps in the reordering queue, the new timer will not be opened until

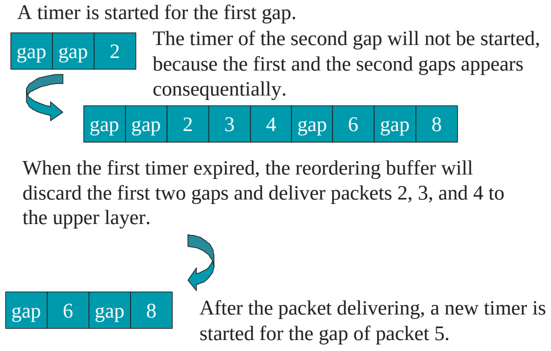

Fig. 4. An example of the timer-based stall avoidance method.

the first timer is expired. In this case, the gap processing time for these two gaps will be $2T_{exp}$ TTIs. Assume that a gap occurs in each slot is equally likely. Thus, within $[T_{exp}, 2T_{exp}]$ the probability of occurring a gap is $\frac{1}{T_{exp}}$. Then, we can compute the average gap processing time for the timer-based stall avoidance mechanism is as follows:

$$\begin{aligned} \bar{T}_{timer} &= \sum_{t=T_{exp}}^{2T_{exp}} \frac{t}{T_{exp}} \cdot t \\ &= \frac{3}{2} T_{exp} \quad . \end{aligned} \tag{1}$$

■

To determine the value of the optimal timer is not an easy task. The timer needs to be long enough for possible retransmission. On the other hand, the timer can not be too long since the timer-based stall avoidance mechanism becomes inefficient.

B. Window-Based Scheme

In [4], a window-based stall avoidance mechanism was proposed to enhance the performance of the timer based shceme. The window-based stall avoidance mechanism can remove the gap in the re-ordering queue before the timer expire. When the reordering buffer is in the overloaded condition.

For example, Figure 5 shows a window-based stall avoidance mechanism with a window size of seven. In this example, the reordering buffer receives packet 2, 3, 4, and 6 but packet 0, 1, and 5. Assume that at this moment, packet 7 arrives. Because the re-ordering buffer is overloaded, the window-based stall avoidance mechanism will start removing the gaps for packet 0 and 1, and send packets 2, 3, 4 to the upper layer.

Proposition 2, we derive the average gap processing time for the window-based stall avoidance mechanism.

Proposition 2 *The average gap prossing time of the window-based stall avoidance mechanism is* $\frac{3}{2}W$*, where* W *is the window size.*

Proof: Assume that all the gaps in the reordering buffer are processed by the window-based scheme before timer expiration. To further facilitate the derivation of the average gap processing time, we first derive the maximum and the minimum gap processing time. Once a gap appears in the reordering buffer, the window-based stall avoidance algorithm has to wait for the reordering buffer to become overloaded before removing the gap from the buffer. Assume the window size is W and denote the probability of successful transmission in each time slot as P_s. Every gap has to wait until W packets are successfully transmitted. It is noted that any gap occurred will be regarded as a successful transmission from the view point of the transmitter because the NACK control signal becomes an ACK signal. Let P_n represent the probability of successful packet transmission at the n^{th} trial. Then, we can write P_n as

$$P_n = (1 - P_s)^{n-1} P_s \ . \tag{2}$$

Note that the maximum number of trials cannot be larger than the timer expiration value. Hence, the average time for successful packet transmission can be expressed as

$$\sum_{n=1}^{N} nP_n = \sum_{n=1}^{N} n(1 - P_s)^{n-1} P_s \ . \tag{3}$$

As mentioned above, the receiver need to wait for W successful transmissions before removing a gap. Therefore, the waiting time can be written as

$$W \sum_{n=1}^{N} nP_n = W \sum_{n=1}^{N} n(1 - P_s)^{n-1} P_s \ . \tag{4}$$

According to (4), we find that the maximum waiting time is equal to $2W$ with $P_s = 0.5$ and the minimum value is W with $P_s = 1$.

Similar to the timer-based scheme, assume the probability of having a gap in the interval $[W, 2W]$ is $\frac{1}{W}$. Then the average gap processing time for the window-based stall avoidance mechanism is the following:

$$\begin{aligned} \bar{T}_{window} &= \sum_{i=W}^{2W} \frac{i}{W} \\ &= \frac{3}{2}W \ . \end{aligned} \tag{5}$$

■

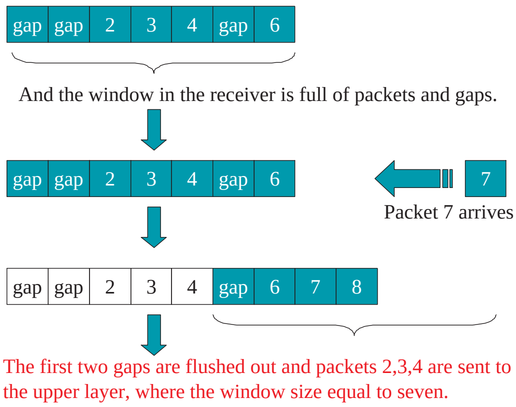

Fig. 5. An example of the window-based stall avoidance method.

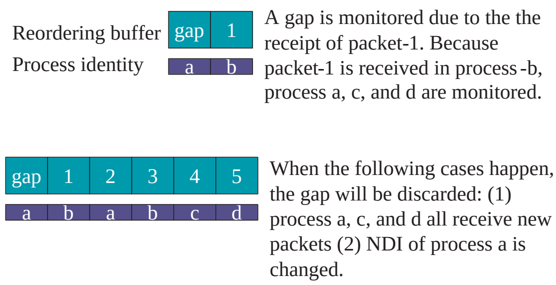

Fig. 6. An example of the indicator-based stall avoidance method.

C. Indicator-Based Scheme

In, the indicator based scheme of [5], [6], the idea of monitoring the activity of each process in the multi-channel SAW H-ARQ was introduced. This idea can expedite the process of forwarding the data in the reordering buffer to the upper layer when a gap occurs in the received packet sequences. For example, Figure 6 shows a scenario of a qual-cahnnel SAW H-ARQ, where process a has a gap and process b is receiving packet 1. At this moment, the indicator-based mechanism start mornitoring the activity of processes a, c, and d. If the receiver finds that the NDIs of process a is indeed sending packet 2 as shown in the figure, the indicator-based stall mechanism will start forwarding the data in the re-ording queue to the upper layer even before the timer expires and the window is overloaded.

In this example, when the new data packet 2 is received at process a, it is implied that packet 0 will no longer be sent by process a in the future. Thus, the concept of monitoring the activity of each H-ARQ process and using new data indicator can enhance the performance of the timer-based and the window-based stall avoidance mechanisms. However, the analysis of the indicator-based scheme is very involved. We will perform simulation to evaluate the performance of the

TABLE I

SIMULATION PARAMETERS.

N-channel SAW H-ARQ	N=4
Frame Size	320 bits
Error Checking	CRC
TTI	2 msec
Timer Expiration	120 TTIs
Sliding window size	32 Packets
Modulation Scheme	BPSK
Doppler frequency	10 or 100 Hz

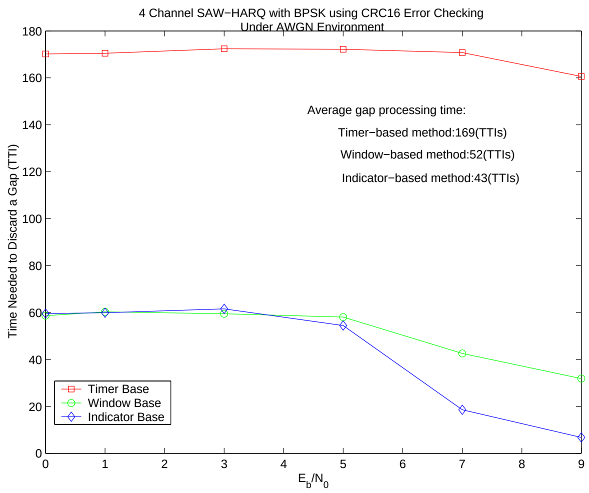

Fig. 7. Performance comparison in terms of gap processing time under AWGN environment.

indicator-based stall avoidance mechanism in Section IV.

IV. SIMULATION RESULTS

We perform simulation to compare the performance of the three stall avoidance mechanism. We observe simulation parameter listed in Table.I.

Figure 7 shows the performance of the gap processing time for the three stall avoidance mechanisms in an AWGN channel. One can consider that for the time required to remove a gap for three algorithms is : $\bar{T}_{timer} > \bar{T}_{window} > \bar{T}_{indicator}$. For example, when $\frac{E_b}{N_0}$ is 7 dB, the indicator-based mechanism only needs 18.55 TTIs to remove the gap, while the window-based and timer-based methods require 42.56 and 170.8 TTIs, respectively. This implies that the indicator algorithm can stop waiting for a retransmission faster than the other two algorithms.

By applying $T_{exp} = 120$ and $W = 32$ to Propositions 1 and 2, we obtain that theoretical values of average gap processing time, which are 180 TTIs and 48 TTIs for the timer-based and the window-based schemes, respectively. Our simulations show that the average gap processing time for the timer based method is 169 TTIS and that for the window-based scheme is 52 TTIs.

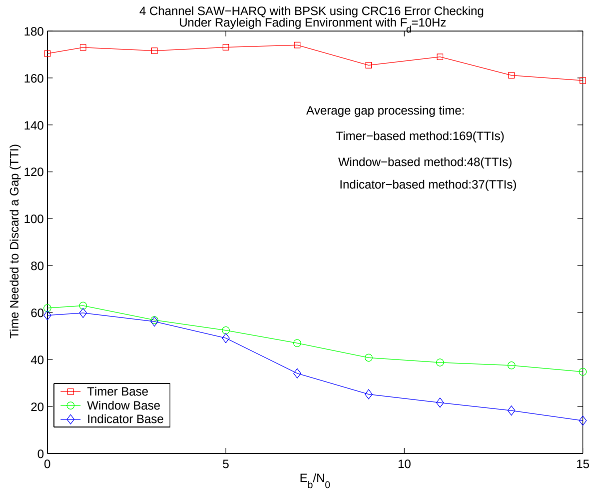

Fig. 8. Performance comparison in terms of gap processing time under Rayleigh fading channel environment with $f_d = 10Hz$.

TABLE II

COMPARISON OF GAP PROCESSING TIME BETWEEN THAT OF $f_d = 10Hz$ AND $f_d = 100Hz$.

	Gap Processing Time (TTI)			
$E_b/N_0(dB)$	Window-Based $f_d = 10Hz$	Indicator-Based $f_d = 10Hz$	Window-Based $f_d = 100Hz$	Indicator-Based $f_d = 100Hz$
5	52.46	49.14	54.35	49.47
7	47.00	34.08	52.24	39.73
9	40.79	25.20	47.52	28.79
11	38.73	21.61	46.48	23.06
13	37.48	18.23	40.51	19.24
15	34.76	13.94	38.08	15.51

Figure 8 shows the performance of three avoidance mechanisms in a Rayleigh fading channel. Compared to the case in an AWGN channel, the gap processing time for both the window-based and indicator-based methods are longer in the Rayleigh fading channel. The effect of Rayleigh fading may increase the retransmission time for a missing packet. Thus, fading may increase gap processing time.

Table I evaluates the impact of Doppler frequency on the gap processing time. From the table, we find that the higher the Doppler frequency, the longer the gap processing time.

V. Conclusion

In this paper, we have investigated the performance of three stall avoidance mechanisms for the WCDMA system with HSDPA in both the AWGN channel and the Rayleigh fading channel. Our results show that the indicator based scheme performs the best. The performance of the windowed based method is worse than the indicator based scheme, but still better than the window based method. The proposed analytical model for estimating the gap processing time of the stall avoidance mechanisms can be useful to provide the guideline in the design of fast retransmission mechanisms for wireless data networks.

References

[1] 3GPP TR 25.950 V4.0.0, "UTRA high speed downlink packet access," Mar. 2001.

[2] 3GPP TSG-RAN WG2 R2-020945, "ACK/NACK power offsets in case of realistic channel estimation," May 2002.

[3] 3GPP TS 25.308 V5.2.0, "High speed downlink packet access (HSDPA) overall description," Mar. 2002.

[4] 3GPP TS 25.321 v5.0.0, "Medium access control (MAC) protocol specification," Mar. 2002.

[5] 3GPP WG2-30 R2-021725, "Stall avoidance schemes in HARQ entity," June 2002.

[6] 3GPP WG2-31 R2-021974, "Stall avoidance schemes in HARQ entity," Aug. 2002.

[7] 3GPP TSG-RAN WG2 R2-A010016, "Dual-channel stop-and-wait HARQ," Edinburg, U.K., Jan 15-19, 2001.

成果報告(二)

高速下行鏈路封包存取的封包排程技術之研究
Packet Scheduling for the WCDMA System with High Speed Downlink Packet Access

計畫編號：NSC 91-2219-E009-016
執行期限：91 年 8 月 1 日至 92 年 7 月 31 日
主持人：王蒞君副教授　國立交通大學電信工程系
計畫參與人員：陳名吉　國立交通大學電信工程系

1 中文摘要

有著高速率下行鏈路封包存取(High Speed Downlink Packet Access, HSDPA) **概念的寬頻分碼多工存取系統是一個很有希望的系統，它採用了適應性調變技術、高效率的排程技術以及混合式的自動重送請求技術使得可以在行動蜂巢式環境中達到高達 10 Mbps 的高速率。在這些技術當中，為了增進這種系統的效能，排程演算法則扮演著重要的角色。一個好的排程演算法目標在於從眾多使用著中，考量到通道的影響、延遲時間之議題以及公平性後選出最適合的使用者出來。在這計劃及**[4]**中，我們採用了一個公平性指標來檢驗現有適合用在此高速率下行鏈路封包存取概念之排程演算法的公平性效能，這些演算法包括最大信號干擾比 (maximum C/I) 排程法、知更鳥式循環(Round Robin)排程法、比例式公平(proportional fair)排程法以及指數型法則(exponential rule)排程法。我們發現現有的排程演算法在此公平性指標上的表現並不是那麼公平。因此，驅使了我們提出一個新的排程演算法，叫做序列式指數型法則(queue-based exponential rule)排程法，來提供比比例式公平排程法以及指數型法則排程法還要好的公平性效能，並且保持高流通量以及低延遲時間的效能。**

關鍵詞：**媒體控制層，高速下行鏈路封包存取**

Abstract

The wideband code division multiple access (WCDMA) system with high speed downlink packet access (HSDPA) is an important next generation wireless system. By adopting adaptive modulation, efficient scheduling, and hybrid automatic repeat request technologies, it can support data rates up to 10 Mbps in the mobile cellular environment. Among these techniques, the scheduling algorithm plays a key role in realizing the HSDPA concept. A good scheduling algorithm should consider all the important factors including channel impact, delay issue, and fairness. In this paper, we adopt a fairness index to examine the fairness performance of current link adaptation based scheduling algorithms, including the maximum C/I, round robin, proportional fair, and exponential rule schedulers. When supporting multi-type services, we find that the fairness performance of current scheduling algorithms, including the round robin scheduler, can be further improved even though the round robin scheduler is viewed as the scheduler of the greatest fairness. Thus, we are motivated to develop a new scheduling algorithm, namely the queue-based exponential rule scheduler. Through simulations, we show that in the context of multi-type services the fairness performance of queue-based exponential rule scheduler can outperform all the schedulers in the time-multiplexing fashion, while maintaining good throughput and delay performance.

Keywords: MAC, Scheduling, CDMA, HSDPA.

[1]**本文已送到** IEEE Wireless Communications and Network Conference (WCNC) 2003 **刊登，詳細內容如附件。**

[1]**本文已送到** IEEE Wireless Communications and Network Conference (WCNC) 2003 **刊登，詳細內容如附件。**

Comparisons of Link Adaptation Based Scheduling Algorithms for the WCDMA System with High Speed Downlink Packet Access

Li-Chun Wang † and Ming-Chi Chen
Department of Communication Engineering
National Chiao Tung University, Hsinchu, Taiwan
lichun@cc.nctu.edu.tw, Tel: +886-3-5712121 ext. 54511

***Abstract*— The wideband code division multiple access (WCDMA) system with high speed downlink packet access (HSDPA) is an important next generation wireless system. By adopting adaptive modulation, efficient scheduling, and hybrid automatic repeat request technologies, it can support data rates up to 10 Mbps in the mobile cellular environment. Among these techniques, the scheduling algorithm plays a key role in realizing the HSDPA concept. A good scheduling algorithm should consider all the important factors including channel impact, delay issue, and fairness. In this paper, we adopt a fairness index to examine the fairness performance of current link adaptation based scheduling algorithms, including the maximum C/I, round robin, proportional fair, and exponential rule schedulers. When supporting multi-type services, we find that the fairness performance of current scheduling algorithms, including the round robin scheduler, can be further improved even though the round robin scheduler is viewed as the scheduler of the greatest fairness. Thus, we are motivated to develop a new scheduling algorithm, namely the queue-based exponential rule scheduler. Through simulations, we show that in the context of multi-type services the fairness performance of queue-based exponential rule scheduler can outperform all the schedulers in the time-multiplexing fashion, while maintaining good throughput and delay performance.**

Index Terms - Scheduling, HSDPA, WCDMA systems.

I. INTRODUCTION

In order to satisfy the fast growing demand of the wireless packet data services, the concept of high speed downlink packet access (HSDPA) is proposed as an evolution for the wideband code division multiple access (WCDMA) system [1]. The goal of the WCDMA system with HSDPA is to support peak data rates from 120 kbps to 10 Mbps by adopting many advanced techniques, such as fast link adaptation, fast physical layer retransmission, and efficient scheduling techniques [2]. A fast link adaptation mechanism can enhance throughput performance by adapting modulation and coding schemes in the rapidly changing radio channel. The hybrid automatic repeat request (HARQ) technique can improve the radio link performance by combining retransmitted packets with previous erroneous packets. Scheduling is the key to achieving fairness in a shared channel for multiple users. Basically, a scheduling algorithm is to select a most suitable user to access the channel in order to optimize throughput, fairness, and delay performances.

Recently, scheduling has attracted much attention for wireless data networks because it can exploit the multi-user diversity [3] [4]. In the traditional voice-oriented cellular network, the fluctuation of the fast fading is viewed as a drawback. However, channel variations can be also beneficial to the wireless *data* network. Because data services can tolerate some delay, a scheduling mechanism can be designed to select the user with the highest channel peak and serve one user at a time in a time-multiplexing fashion. With a higher channel peak, a more efficient modulation/coding scheme can be applied to enhance data rates. In general, a larger dynamic range of channel variations can yield higher channel peaks, thereby delivering larger multi-user diversity gain.

In addition to throughput optimization, service delay and fairness are another two important aspects needed to be taken into account in designing a good scheduling algorithm. The reason why wireless scheduling can improve system throughput lies in the fact that data services can tolerate certain level of delay. Nevertheless, a constraint on the maximum service delay is still necessary. In wireless systems, mobile users are located at different locations with different channel conditions. If a scheduler always selects the user with the best channel condition, some users at the cell boundary, for instance, may never have a chance to access the system. Hence, how to design a scheduling algorithm to achieve high throughput subject to delay and fairness constraint becomes a crucial and challenging issue for the wireless data network.

In the literature, wireless scheduling algorithms can be categorized into two major types according to the considered channel models. First, the wireless scheduling algorithms in [5] [6] [7] considered a two-state on-off Markov channel model. Because of simplicity, the two-state Markov channel model is suitable to examine the fairness performance of scheduling algorithms. However, using the simple two-state Markov channel has limitations in capturing actual radio channel characteristics. Second, some wireless scheduling algorithms like [1] [8] [9] [11] considered a more practical radio channel model with the emphasis on exploiting the multi-user diversity. In [1], the maximum carrier to interference ratio (C/I) scheduler is designed to assign the channel to the user

[1]Contact author.

[2]This work was supported by the National Science Council and the MOE program for Promoting Academic Excellence of University under the grant number. 91-2219-E-009-016, EX-91-E-FA06-4-4 and 89-E-FA06-2-4.

with the best C/I. Obviously, the maximum C/I scheduler fully utilizes the multi-user diversity, but is an unfair scheduling policy. In contrast to the maximum C/I scheduler, a fair time scheduler (or called round robin scheduling algorithm [8]) allocates the channel to users in sequence with equal service time. Clearly, the fair time scheduler does not consider the channel effect. In [2] [8] [9] [10], the proportional fair scheduler was proposed for the IS-856 system and the WCDMA system. The proportional fair scheduler improves the fairness performance of the maximum C/I scheduler at the cost of lowering system throughput. However, in [11] it was pointed out that the proportional fair scheduling algorithm does not account for service delays. Thus, the authors in [11] proposed the exponential rule scheduler to improve the delay performance of the proportional fair scheduler. It was proved that the exponential rule scheduler is throughput optimal in the sense of making a service queue stable [12]. In [13], it was concluded that the exponential rule scheduler is superior to the proportional fair scheduler since it provides excellent latency performance even with a slightly lower system throughput. Nevertheless, the factor of queue length is still not explicitly considered in the exponential rule scheduling policy.

To our knowledge, a wireless scheduling policy with consideration of all the factors of channel variations, service delay, and queue length is still lacking in the literature. Consequently, we are motivated to develop such a wireless scheduling algorithm. The contributions of this work are two folds. The first contribution of this work is to propose a queue-based exponential rule scheduler to explicitly take account of all the factors consisting of queue length, service delay, and channel variations. We find that the queue-based exponential rule scheduler can further improve the fairness performance as compared to the original exponential rule scheduler, while maintaining the same throughput and delay performances.

Second, in the context of multi-type services, we suggest a fairness index in evaluating the fairness performance of the link adaptation based wireless scheduling algorithms, i.e., the maximum C/I, proportional fair, and exponential rule schedulers. This fairness index is modified from the one used in the two-state Markov channel based wireless scheduling algorithms [5] [6] [7] [14]. Interestingly, we find the fairness of current link adaptation based wireless scheduling algorithms [1] [8] [9] [11] is not clearly specified. Thus, we are motivated to adopt a formal fairness index to evaluate the fairness performance of these link adaptation based wireless scheduling algorithms. Using this fairness index to compare scheduling algorithms is important, especially in the case of the multi-type services. For example, the fair time (or round robin) scheduler is viewed as the performance upper bound in terms of fairness for most link adaptation based wireless scheduling algorithms. However, this upper bound is only valid under the assumption that all users subscribe the same type of service in any particular duration. In the case of multi-type services, the fair time scheduler can only guarantee the equal access time for multiple users, but can not ensure to satisfy the different requirements for different users. Thus, even the fair time scheduler may not be the fairest scheduling algorithm in supporting the multi-type services. Thus, to support multi-type services, it is important to re-evaluate the fairness of these link adaptation based wireless scheduling algorithms based on a formal fairness index. Our simulation results show that in the time-multiplexing fashion, the fairness performance of the exponential rule scheduler is very close to that of the fair time scheduler and both schemes are superior to the proportional fair scheduler. On the other hand, in the code-multiplexing fashion, we find that the exponential rule scheduler is only slightly better than the proportional fair rule, and both schemes are worse than the fair time scheduler.

The rest of this paper is organized as follows. Section II briefly introduces the background of the HSDPA concept in the WCDMA system. Section III describes existing link adaptation based scheduling algorithms. We discuss our new proposed queue-based exponential rule scheduling algorithm in Section IV. Simulation results are presented in Section V. Finally, we give our concluding remarks in Section VI.

II. High Speed Downlink Packet Access

In this section, we introduce three key technologies in implementing the HSDPA concept, including adaptive modulation and coding, fast scheduling, and multi-code assignment.

A. Adaptive Modulation and Coding

Adaptive modulation and coding is a link adaptation technique to adapt transmission parameters to the time varying channel conditions. The basic principle in applying adaptive modulation and coding is to assign higher order modulation and higher code rate to the users in favorable channel conditions, whereas allocate lower order modulation and lower code rate to the users in unfavorable channel conditions. To obtain the channel conditions, the receiver is required to measure the channel conditions and feedback to the the transmitter periodically, e.g., 2 msec of every transmission time interval (TTI) in HSDPA. The major benefit of adaptive modulation and coding is to deliver higher data rate to the user with better channel conditions, thereby increasing the average throughput of the cell. Table I lists the modulation and coding schemes used in HSDPA.

B. Fair Scheduling

In the downlink shared channel, scheduling techniques can be used to exploit multi-user diversity by taking advantage of short-term channel variations of each user terminal. With scheduling the selected user is always served in the high channel peak or a constructive fade. When combined with adaptive modulation and coding, the scheduled user can therefore have a better chance to transmit at a higher rate with higher order modulation and higher code rate.

C. Multi-Code Assignment

To achieve a higher data rate in HSDPA, adaptive modulation and coding technique needs to function together with multi-code assignment. Because the number of modulation and coding schemes is limited (e.g., only 5 formats in HSDPA), the dynamic range of selecting adaptive modulation and coding scheme may not be enough, e.g., 15 dB of dynamic range in an example of Fig. 51 in [1]. To increase the dynamic range of adapting transmission parameters to channel variations, selecting both the modulation/coding scheme and the number of codes is an effective method. Figure 1 shows the $hull$ curve of combining 5 modulation/coding schemes with the multi-code operation of 10 codes. As shown in the figure, the dynamic range of selecting transmission parameters for different E_c/I_{oc} (i.e., signal to interference ratio (SIR)) becomes -5 to 20 dB, which is 10 dB larger than the single code case of [1]. Generally speaking, it is better to increase the number of multi-code first before transiting to the next higher order of modulation/coding scheme based on a spectral efficiency viewpoint. Note that Fig. 1 in this paper is produced by referring to Fig. 51 of [1], both of which do not consider HARQ. If ARQ is considered, we can straightforwardly use the hull curve of adaptive modulation and coding with ARQ in the single code operation (e.g., Fig. 50 of [1]) to produce a hull curve of the adaptive modulation/coding scheme with HARQ for the multi-code operation.

III. Current Link Adaptation Based Scheduling Algorithms

As mentioned, the key factor in determining the performance of the WCDMA system with HSDPA is the link adaptation based scheduling algorithm [1] [8] [9] [11]. We will evaluate these scheduling algorithms in terms of three performance metrics: system throughput, delay, and fairness. In order to achieve the multi-user diversity, a good scheduling algorithm should take account of channel variations. The more the multi-user diversity, the higher the system throughput. However, for some delay-sensitive services, such as streaming video, a good scheduler also needs to pay attention to the latency performance. Furthermore, from an individual user standpoint, it is desirable to receive the service as fair as others. Thus, to design a good scheduling algorithm, we need to make a better tradeoff design between these three performance metrics.

Before we detail the scheduling algorithms, let us first define the following notations:

- k: the index of the k^{th} transmission time interval (TTI).
- $\gamma_i(k)$: the short-term SIR of user i averaged in the $(k-1)^{th}$ TTI.
- $\overline{\gamma_i(k)}$: the long-term average SIR of user i observed in $[(k-T),k]$, where T is the length of sliding window in terms of the number of TTIs.
- $d_i(k)$: the delay time for the packet waiting in the head of line (HOL) TTI before getting the service for user i.
- $q_i(k)$: the queue length of user i at the beginning of the k^{th} TTI.

A. Maximum C/I Scheduler

The maximum C/I scheduler always selects the user with the best carrier-to-interference ratio (C/I). At the beginning of each TTI, the scheduler compares the C/I levels of all active users and grants the channel access to the user with the highest C/I level. Specifically, the maximum C/I scheduler will select the user j in the k-th TTI if

$$j = \arg\{\max_i \gamma_i(k)\}. \tag{1}$$

Obviously, the maximum C/I scheduler is the upper bound in terms of system throughput because it can achieve the maximum multi-user diversity. However, those who are located far from the base station may feel unfair owing to the poor radio link condition. This unfairness phenomenon makes the maximum C/I scheduler in practical.

B. Fair Time Scheduler

At any scheduling instant, the fair time scheduler serves all non-empty source queues in a round-robin fashion. That is, the fair time scheduler will schedule user j in the k-th TTI if

$$j = \textit{mod}\,((k-1), N) + 1, \tag{2}$$

where $mod(\cdot)$ denotes the modulus operator and the N is the number of active users in the system. One can easily finds that this fair time scheduling algorithm is independent of the channel characteristics and thus does not exploit the multi-user diversity at all. However, this scheduler has the best delay performance and is much fairer than the maximum C/I scheduler. Note that the fairness here is defined from the viewpoint of equal access probability. In this paper, we will adopt another fairness index used in the two-state Markov channel based wireless scheduling algorithms [14] to examine the fairness performance of the scheduling algorithms.

C. Proportional Fair Scheduler

The proportional fair algorithm was proposed to the HDR system [9]. The basic idea of this algorithm is to select a scheduled user based on the ratio of the short-term SIR over the long-term averaged SIR value with respect to each active user. The proportional fair scheduler takes advantage of the temporal variations of the channel to increase system throughput, while maintaining certain fairness among all active users. Specifically, the proportional fair scheduler will schedule user j in the k-th TTI if

$$j = \arg\{\max_i \frac{\gamma_i(k)}{\overline{\gamma_i(k)}}\}. \tag{3}$$

Based on the above criterion, the $\overline{\gamma_i(k)}$ is the average SIR measured over a sliding window as follows,

$$\overline{\gamma_i(k+1)} = \begin{cases} (1-\frac{1}{T})\overline{\gamma_i(k)} + \frac{1}{T}\gamma_i(k) \\ \text{if user } i \text{ is scheduled,} \\ (1-\frac{1}{T})\overline{\gamma_i(k)} \\ \text{if user } i \text{ is not scheduled.} \end{cases}$$

Note that the proportional fair scheduler does not consider the delay issue in each user's service queue and thus has poor delay performance.

D. Exponential Rule Scheduler

The exponential rule is a modified version of proportional fair. This scheduler further takes delay issue into consideration [11] [13]. This policy tries to balance the weighted delay of all active users when the differences of weighted queue delay among users become significant. For the k^{th} TTI, the exponential rule scheduler will choose user j if

$$j = \arg\left\{\max_i a_i \frac{\gamma_i(k)}{\overline{\gamma_i(k)}} \exp\left(\frac{a_i d_i(k) - \overline{ad(k)}}{1+\sqrt{\overline{ad(k)}}}\right)\right\}, \quad (4)$$

where

$$\overline{ad(k)} = \frac{1}{N}\sum_{i=1}^{N} a_i d_i(k), \quad (5)$$

and $a_i > 0,\ i = 1, ..., N$ are selected weights to characterize the desired quality of service.

To obtain the intuition behind the exponential rule scheduler, let us focus on the exponent term of (4). If an active user has a larger weighted delay than others by more than $\sqrt{\overline{ad(t)}}$, then the exponent term will become dominant and even exceed the impact of channel effect. In such a case, this user will get higher priority of the access opportunity. On the other hand, for small differences in the weighted delay, the exponent term will approach to 1 and the exponential rule scheduler is exactly the same as the original proportional fair rule scheduler. Note that the factor 1 in the denominator of (4) is to prevent the exponent term from increasing dramatically when $\sqrt{\overline{ad(t)}}$ is too small. In brief, the exponential rule scheduler incorporates both the effect of channel variations and service delay and aims to further improve the performance over the proportional fair scheduler.

IV. Proposed Queue-Based Exponential Rule Scheduler

In this section, we will first define the fairness and then propose an improved version of exponential rule scheduler.

A. Fairness Definition

Based on the traditional fluid fair queueing in wireline networks, backlogged flows are served in proportion to their weighted rate. Mathematically, for any time interval $[t_1, t_2]$, the channel capacity allocated to flow i, denoted as $W_i(t_1, t_2)$, should satisfy the following condition [14] [15]:

$$\left|\frac{W_i(t_1,t_2)}{r_i} - \frac{W_j(t_1,t_2)}{r_j}\right| = 0,\ \forall i, j \in B(t_1, t_2), \quad (6)$$

where r_i and r_j are the weights of flows i and j, respectively, and $B(t_1, t_2)$ is the set of backlogged flows during (t_1, t_2).

However, this condition in terms of bits can not be maintained in a practical packet switched network. The goal of the packetized fair queueing algorithm is to minimize the difference of $|W_i(t_1, t_2)/r_i - W_j(t_1, t_2)/r_j|$. Therefore, we define a fairness index as follows,

$$FI = \frac{1}{l}\left|\frac{W_i(t_1,t_2)}{r_i} - \frac{W_j(t_1,t_2)}{r_j}\right|, \quad (7)$$

where l is a normalization factor of the packet size. In wireless networks, many wireless scheduling algorithms, such as IWFQ [5], CIF-Q [6], and WFS [7], were proposed to minimize the value of FI during a long period, i.e., achieving the long-term fairness. Here, we define the packet size as the number of data bits that can be transmitted at the minimum data rate during one transmission time interval (TTI).

B. Queue-Based Exponential Rule Scheduler

The exponential rule scheduler considers the effect of the delay time in the head of line (HOL) TTI. However, it does not *explicitly* incorporate the factor of queue length into the scheduling policy. By observing (7), we note that to achieve fairness not only the HOL delay is required to be considered, but also the queue length. Consequently, we are motivated to propose a queue-based exponential rule scheduler as follows. In the k^{th} TTI, the proposed scheduler will choose user j if

$$\begin{aligned} j \quad &= \arg\{\max_i a_i \tfrac{\gamma_i(k)}{\overline{\gamma_i(k)}} \exp\left(\tfrac{a_i d_i(k) - \overline{ad(k)}}{1+\sqrt{\overline{ad(k)}}}\right) \cdot \\ &\exp\left(\tfrac{q_i(k)-\overline{q(k)}}{1+\overline{q(k)}}\right)\}, \end{aligned} \quad (8)$$

where

$$\overline{ad(k)} = \frac{1}{N}\sum_{i=1}^{N} a_i d_i(k), \quad (9)$$

and

$$\overline{q(k)} = \frac{1}{N}\sum_{i=1}^{N} q_i(k). \quad (10)$$

The basic idea of second exponent term in (8) is to balance the service queue length among multiple users. Moreover, in order to prevent the second exponent term from exceeding that of first exponent term, the denominator of the second exponent term does not take the square root as that of the first exponent term.

C. Time-Multiplexing Fashion v.s. Code-Multiplexing Fashion

In HSDPA, there are five available modulation and coding schemes and sixteen orthogonal variable spreading factor (OVSF) codes [8]. There are two methods to implement the multi-code operation. One is to assign all available multi-codes to one user at a time in a pure time-multiplexing fashion. The other is to assign different active users to each code in a code-multiplexing fashion. For the code-multiplexing multi-code assignment, we can treat each code as a server and scheduling algorithms will select suitable users for multiple servers simultaneously. Note that the total power budget in each base station is equally shared by all codes for both methods. The performance comparison of these two methods will be discussed in Section V.

V. Simulation Results

Through simulation, we compare the performances of different scheduling algorithms. Three types of simulation configurations are considered. First, we focus on the performance with the single code operation in the time-multiplexing fashion. In this case, the maximum data rate for each scheduled mobile user is 720 kbps. In the second simulation configuration, we focus on the performance of the multi-code operation in the time-multiplexing fashion. With maximum 10 codes, the maximum achieved data rate in this configuration becomes 7.2 Mbps. For the third simulation configuration, we investigate the performance of the multi-code operation in the code-multiplexing fashion.

A. Simulation Model

We consider a cell layout with a center cell surrounded by other six neighboring cells. We only focus on the performance of the center cell and treat other cells as the sources causing the downlink interference. The mobile users are uniformly located in the center cell. We assume that 60% of the total base station transmitted power is allocated in supporting the HSDPA services. In HSDPA, there are 16 orthogonal codes with processing gain of 12 dB. Nevertheless, many other dedicated, shared and common channels may also need to use some codes. Thus, for the multi-code operation in HSDPA, the suggested maximum number of codes can be assigned to a user is 10 [8]. Table II lists other parameters used in our simulations [1].

We apply the $hull$ curve of the link level simulation results obtained from [1] into our system level simulation. For each randomly located mobile user, the system will first calculate the corresponding received signal to interference ratio (SIR). It is assumed that the measured SIR in each mobile can be correctly sent back to the base station in time. We consider a mobile at a speed of 3 km/hr in a flat Rayleigh fading channel, or equivalently the maximum Doppler frequency f_d = 5.5 Hz.

B. Single-Code Operation in the Time-Multiplexing Fashion

Figure 2 compares the fairness performance according to the fairness definition of (7) for all the considered schedulers. Based on (7), the smaller the fairness index, the fairer the system. Obviously, the scheduler of the least fairness is the maximum C/I scheduler. The fair time scheduler, proportional fair scheduler and exponential rule scheduler almost have the same fairness performance. It is noteworthy that the proposed queue-based exponential rule scheduler is fairer than all the other schedulers. Specifically, the proposed new scheduling algorithm improves the fairness index values by 31% to 50% as compared to the original exponential rule scheduler.

Figure 3 compares the throughput performance of the proposed scheduling algorithm with four existing scheduling algorithms. For the maximum C/I scheduler, one can find that when the number of users in the system increases, the system throughput improves because more multi-user diversity gain are achieved. The fair time scheduler results in the lowest throughput, whereas the throughput of the proportional fair scheduler is between the maximum C/I scheduler and the fair time scheduler. Our proposed queue-based exponential rule scheduler performs the same as the original exponential rule scheduler, but has lower throughput than the proportional fair scheduler. Note that when the number of users per cell increases, the achieved average throughput of both the exponential rule scheduler and queue-based exponential rule scheduler slightly decrease because the effect of the exponent term becomes dominant for the case with a large number of users in the system. Although the proportional fair scheduler has better throughput, the exponential rule scheduler has better delay performance, which will be discussed next.

Figure 4 shows the delay performance of the considered schedulers in terms of the maximum delay for the packet waiting in the HOL TTI. In general, higher traffic load causes longer delay because of more contenders. One can find that both the maximum C/I and proportional fair schedulers have very poor delay performance. As expected, the fair time scheduler performs the best. Note that both the proposed queue-based exponential rule scheduler and the original exponential rule scheduler perform well and have the same delay performance, while the proposed queue-based exponential rule scheduler has better fairness performance as shown in Fig. 2 previously.

C. Multi-Code Operation in the Time-Multiplexing Fashion

Figures 5, 6 and 7 show the performances of fairness, system throughput and delay for different schedulers with multi-code operation in the time-multiplexing fashion, respectively. Basically, the trends of these performance curves are the same as in the single code case except that for the multi-code operation, the maximum achievable system throughput becomes 7.2 Mbps. Again, the fairness index of the new proposed algorithm improves ranged from 15% to 30% as compared to original exponential rule scheduler, while maintaining the same throughput and delay performance.

D. Multi-Code Operation in the Code-Multiplexing Fashion

Figure 8 shows the fairness performance of all schedulers in the case of code-multiplexing model. Interestingly, we find that the exponential rule scheduler is only slightly better than the proportional fair rule, and both schemes are worse than the fair time scheduler. Nevertheless, when we further evaluate the variance of the fairness index for different schedulers. As shown in Fig. 9, the variations of the fairness index of the queue-based exponential rule scheduler is much smaller than that of the fair time scheduler. This phenomenon implies that the proposed queue-based exponential can still have a high possibility to be fairer than the fair time scheduler even in the code-multiplexing case. In addition, the new proposed algorithm improved the fairness index value by 5% to 15% as compared to the original exponential rule scheduler.

Figure 10 shows the system throughput performance of all schedulers with the multi-code operation in the code-multiplexing fashion. We consider more than 10 users in this simulation. One can find that the code-multiplexing fashion cannot fully take advantage of the multi-user diversity even for the maximum C/I method. This observation shows that assigning only one user at a time can achieve the maximum system throughput, which was also mentioned in [4]. Specifically, comparing Fig. 6 with Fig. 10 in the case of the system load equal to 15 users, the proposed scheduling algorithm can provide 2.3 Mbps of system throughput in the time-multiplexing scheme, while in the code-multiplexing scheme the throughput decreases to 1.8 Mbps.

Figure 11 compares the corresponding delay performance of all schedulers in the code-multiplexing case. Because of more servers during each scheduling instant, the delay performance in the code-multiplexing is much better than that in the time-multiplexing fashion. For example, in the case of 30 users, the maximum delay of the queue-based exponential rule scheduler in the code-multiplexing fashion is reduced to 80 TTIs as compared to 100 TTIs in the time-multiplexing fashion. Thus, when comparing the performance between the second and third simulation configurations, we can conclude that the delay performance of the code-multiplexing scheme is better than the time-multiplexing at the expense of lower system throughput compared to the time-multiplexing scheme.

VI. Conclusion

In this paper, we adopt a fairness index to examine the fairness performance of current link adaptation based scheduling algorithms, including the maximum C/I, fair time, proportional fair, and exponential rule schedulers. Moreover, we have proposed a new queue-based exponential rule scheduler for the HSDPA system. As summarized in Table III, the proposed scheduling algorithm outperforms all the existing scheduling algorithms in terms of fairness for the time-multiplexing fashion. In the time-multiplexing case, the proposed queue-based exponential rule scheduler improves the fairness index in a range of 15% to 50% compared to the original exponential rule algorithm, while in the code-multiplexing case the improvement reduces to the range of 5% to 15%. For the code-multiplexing configuration, Table IV compares the schedulers considered with respect to fairness, delay, and throughput. Note that although the fairness performance of the proposed queue-based exponential rule scheduler is slightly worse than the fair time scheduler in terms of the average fairness index value, the variations of the fairness index value of the proposed scheduler is smaller than the fair time scheduler, thereby still having possibility to be fairer than the fair time scheduler. It is noteworthy that in both the time-multiplexing and code-multiplexing cases, the proposed queue-based exponential rule scheduler and the original exponential scheduler improve the fairness and delay performances over the proportional fair scheduler at the cost of slightly degrading throughput. Of the comparison between code-multiplexing scheduling and time-multiplexing scheduling, we find that the time-multiplexing method is a better choice by taking all the factors of system throughput, service delay, and fairness into consideration.

TABLE I

Modulation and Coding Schemes in the HSDPA Concept

Modulation and coding schemes (MCS)	Modulation	Effective code rate
MCS 1	QPSK	1/4
MCS 2	QPSK	1/2
MCS 3	QPSK	3/4
MCS 4	16-QAM	1/2
MCS 5	16-QAM	3/4

References

[1] 3GPP TR 25.848 V4.0.0, Physical layer aspects of UTRA High Speed Downlink Packet Access, Release 4, Mar. 2001.

[2] T. E. Kolding, F. Frederiksen, P.E. Mogensen, "Performance aspects of WCDMA systems with high speed downlink packet access (HSDPA)," in *Proceedings of 56th IEEE Vehicular Technology Conference* (VTC 2002-fall), vol. 1, pp. 477 -481, Sep. 2002.

[3] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels part I: Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2796 -2815, Nov. 1998.

[4] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," *IEEE International Conference on Communications*, vol. 1,pp. 331-335, Jun. 1995.

[5] S. Lu, V. Bharghavan, and R. Srikant, "Fair queuing in wireless packet networks," in *Proceedings of ACM SIGCOMM '97*, pp. 63-74, Cannes France, Sep. 1997.

[6] T. S. Ng, I. Stoica, and H. Zhang, "Packet fair queueing algorithms for wireless networks with location-dependent errors," in *Proceedings of IEEE INFCOM*, pp. 1103-1111, San Francisco, CA, Mar. 1998.

[7] S. Lu, T. Nandagopal, and V. Bharghavan, "Fair scheduling in wireless packet networks," in *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking*, pp. 10-20, Oct. 1998.

[8] Y. Ofuji, A. Morimoto, S. Abeta, and M. Sawahashi, "Comparison of packet scheduling algorithms focusing on user throughput in high speed downlink packet access," , *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 3, pp. 1462-1466, Sep. 2002.

[9] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR: a high efficiency-high data rate personal communication wireless system," *Proceedings of the IEEE VTC 2000-Spring*, Tokyo, Japan, pp. 1854-1858, May. 2000.

[10] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and S. Viterbi, "CDMA/HDR: a bandwidth efficient high speed wireless data service for nomadic users," IEEE Communications Magazine, vol. 38, no. 7, pp. 70-77, Jul. 2000.

[11] S. Shakkottai and A. L. Stolyar, "Scheduling algorithms for a mixture of real-time and non-real-time data in HDR," Bell Laboratories Technical Report, Oct. 2000.

[12] S. Shakkottai and A. L. Stolyar, "Scheduling for Multiple Flows Sharing a Time-Varying Channel: The Exponential Rule," Bell Laboratories Technical Report, Oct. 2000.

[13] A. Ekşim and M. O. Sunay, "On scheduling for delay tolerant traffic in HDR," IEEE International Symposium on Advances in Wireless Communications, pp. 189-190, Sep. 2002.

[14] I. Stojmenovic, *Handbook of Wireless Networks and Mobile Computing*. 1st ed. John Wiley & Sons, 2002.

[15] V. Bharghavan, L. Songwu, and T. Nandagopal, "Fair queuing in wireless networks: issues and approaches," *IEEE Personal Communications*, vol. 6, no. 1, pp. 44-53, Feb. 1999.

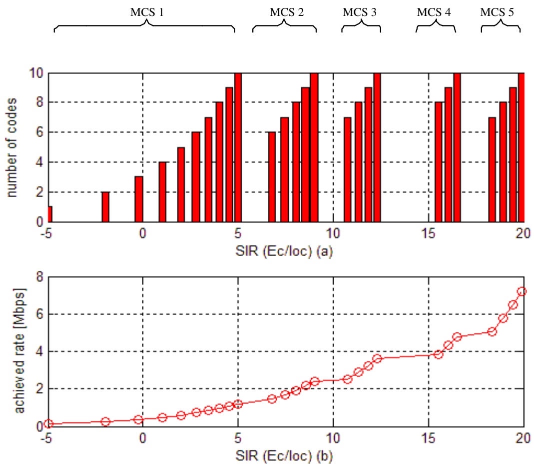

Fig. 1. The *hull* curve with five modulation and coding schemes and the multi-code operation of maximum 10 codes [1].

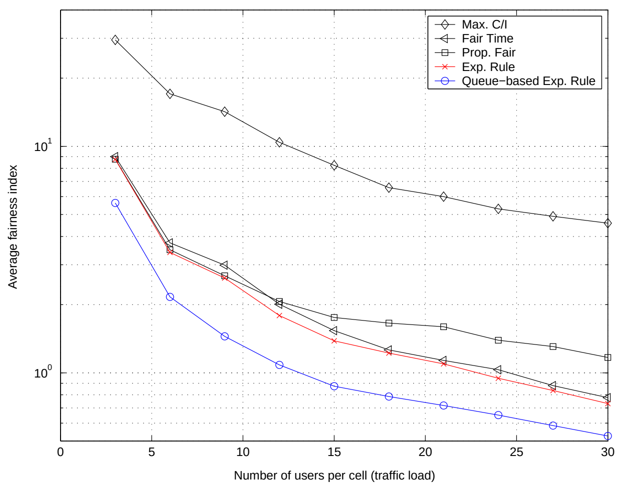

Fig. 2. The average of fairness index value for all schedulers without multi-code operation in a time-multiplexing fashion.

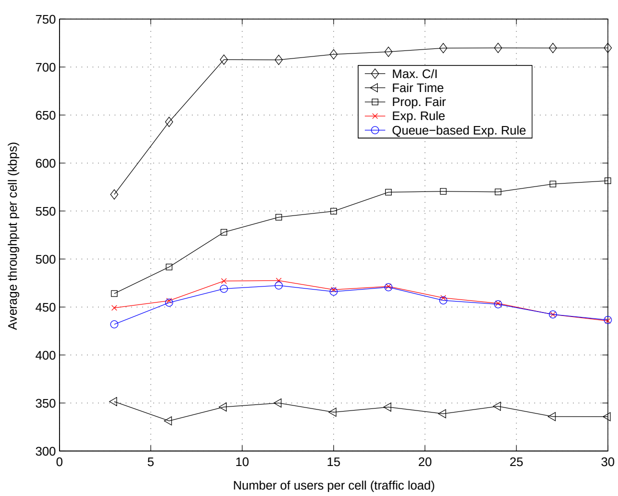

Fig. 3. System throughput performance of all schedulers without multi-code operation in a time-multiplexing fashion.

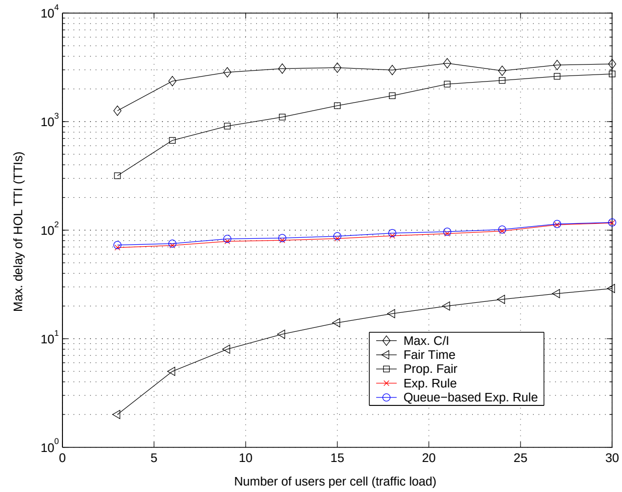

Fig. 4. The performance of maximum delay of all schedulers without multi-code operation in a time-multiplexing fashion.

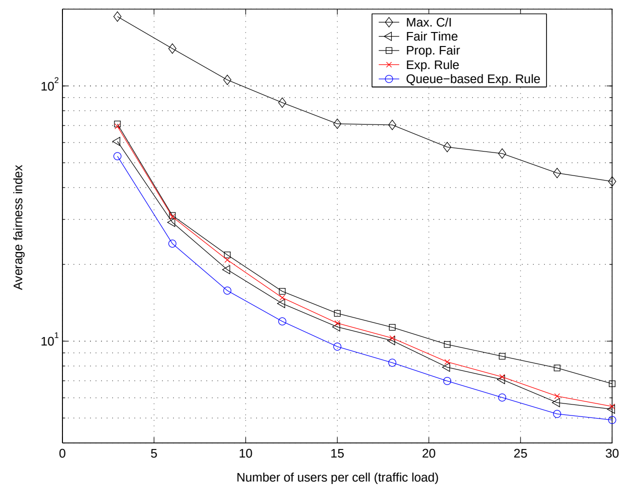

Fig. 5. The average of fairness index value for all schedulers with multi-code operation in a time-multiplexing fashion.

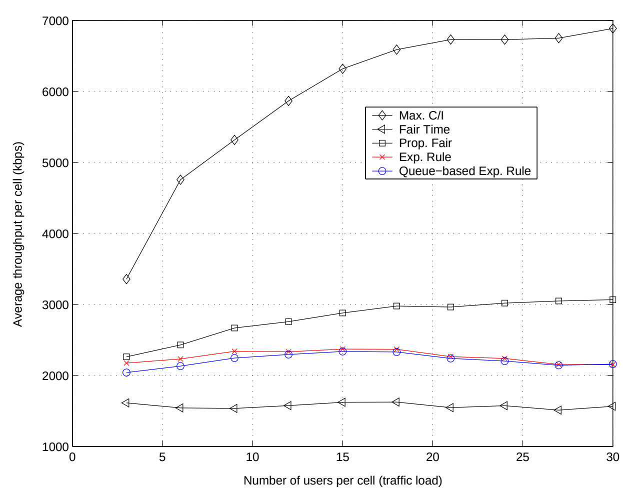

Fig. 6. System throughput performance of all schedulers with multi-code operation in a time-multiplexing fashion.

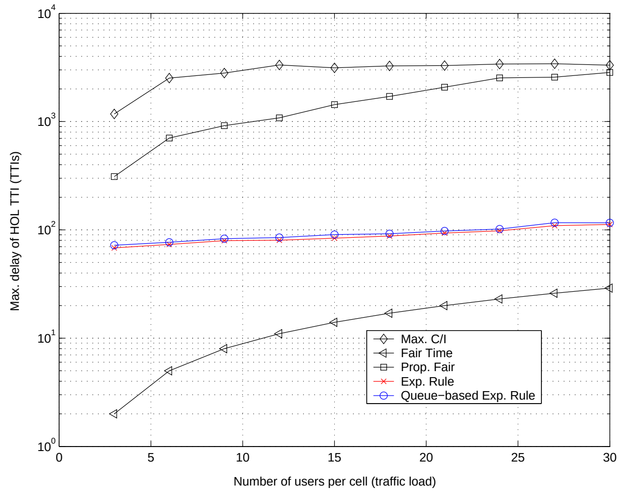

Fig. 7. The performance of maximum delay of all schedulers with multi-code operation in a time-multiplexing fashion.

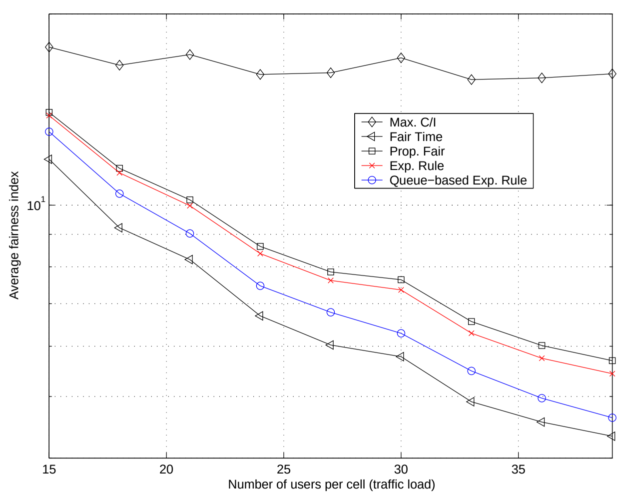

Fig. 8. The average of fairness index value for all schedulers with multi-code operation in a code-multiplexing fashion.

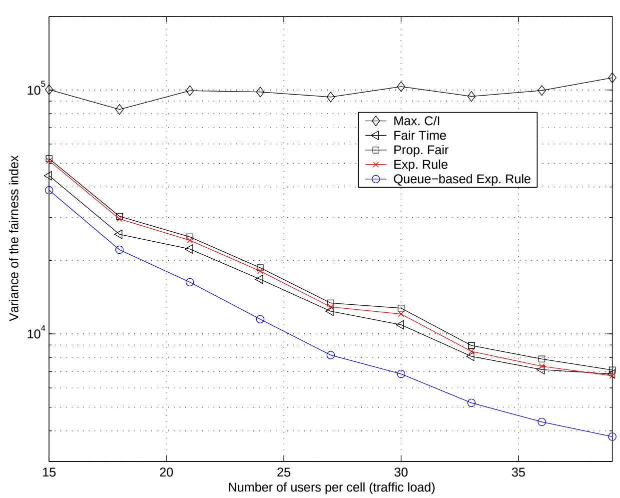

Fig. 9. The variance of fairness index value for all schedulers with multi-code operation in a code-multiplexing fashion.

TABLE II
SIMULATION PARAMETERS

Parameters	Explanation/Assumption
Cell layout	7 hexagonal cells
Cell radius	1.6 km
User location	Uniform distribution
Antenna pattern	Omni-direction
Propagation model	L = 128.1 + 37.6 Log10(R)
HS-DSCH power budget/the total Node-B power	60 %
Shadowing Std. deviation	8 dB
Correlation distance of shadowing fading	50 m
Carrier frequency	2 GHz
Base station total transmit power	44 dBm
Fast fading model	Jakes spectrum
Number of HS-DSCH multi-codes	10
Transmission time interval (TTI)	2 msec
Simulation duration	5000 TTIs

TABLE III
COMPARISON OF LINK ADAPTATION BASED SCHEDULING ALGORITHMS IN TERMS OF FAIRNESS, THROUGHPUT, AND DELAY IN THE TIME-MULTIPLEXING FASHION

Scheduling policy	Fairness	Delay	Throughput
Maximum C/I	Poor	Poor	Excellent
Fair time	Good	Excellent	Poor
Proportional fair	Acceptable	Poor	Good
Exponential Rule	Good	Good	Acceptable
Queue-based exponential Rule	Excellent	Good	Acceptable

p.s. comparison level: excellent $>$ good $>$ acceptable $>$ poor.

TABLE IV
COMPARISON OF LINK ADAPTATION BASED SCHEDULING ALGORITHMS IN TERMS OF FAIRNESS, THROUGHPUT, AND DELAY IN THE CODE-MULTIPLEXING FASHION

Scheduling policy	Fairness	Delay	Throughput
Maximum C/I	Poor	Poor	Excellent
Fair time	Excellent	Excellent	Poor
Proportional fair	Acceptable	Poor	Acceptable
Exponential Rule	Acceptable	Good	Acceptable
Queue-based exponential Rule	Good	Good	Acceptable

p.s. comparison level: excellent $>$ good $>$ acceptable $>$ poor.

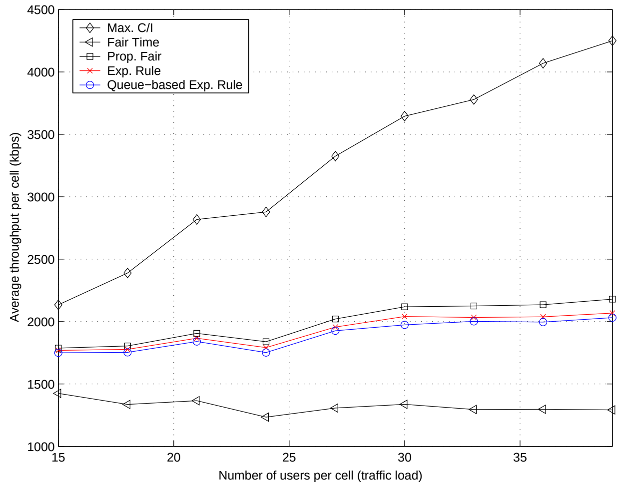

Fig. 10. System throughput performance of all schedulers with multi-code operation in a code-multiplexing fashion.

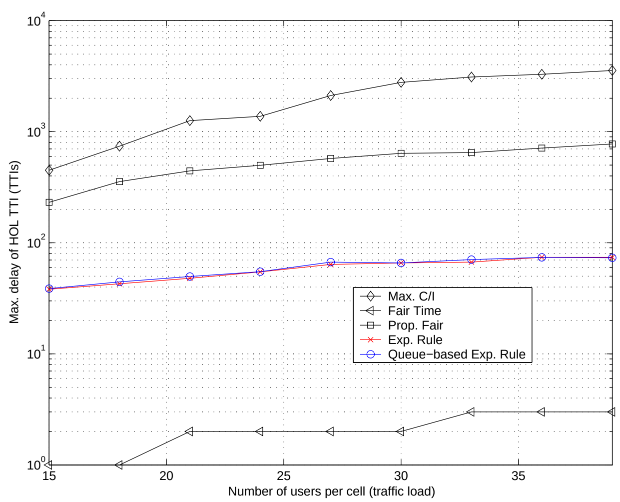

Fig. 11. The performance of maximum delay of all schedulers with multi-code operation in a code-multiplexing fashion.

成果報告(三)

使用鄰近資訊與預測且適用在多媒體環境及分散式的無線區域網路通訊協定 [1]

NICE - A Decentralized Medium Access Control Using Neighborhood Information Classification and Estimation for Multimedia Applications in Ad Hoc 802.11 Wireless

計畫編號：NSC 91-2219-E009-016

執行期限：91 年 8 月 1 日至 92 年 7 月 31 日

主持人：王蒞君副教授　 國立交通大學電信工程系

計畫參與人員：陳銘賓　國立交通大學電信工程系

1 中文摘要

行動分散式區域網路(MANET: mobile ad hoc network)**中媒體存取控制協定的性質須俱備**(1)**符合即時節點要求的服務品質**(2)**分散式**(3)**以流量及能量消耗的觀點達到公平性**(4)**對隱藏節點問題具有免役力。雖然已有很多研究對於**IEEE 802.11 **無線區域網路協定有深入的探討，但卻很少有同時滿足上述四項要求的協定。我們所提出來的協定能夠滿足即時訊務並滿足服務品質的要求且滿足非即時節點的公平性原則。再者我們的提案無需任何集中式的控管，便能夠輕易的在** MANET **上建立。我們針對所提出的協定的流量建立了一個分析模型。最後根據分析模型以及模擬，比較了各協定的流量。**

關鍵詞：**IEEE 802.11， 無線區域網路，媒體控制層，服務品質，分散式無線區域網路**

Abstract

The desired properties of a medium access control (MAC) protocol in mobile ad hoc network (MANET) include (1) meet quality of service (QoS) requirements for real-time nodes, (2) be decentralized, (3) achieve fairness from viewpoint of throughput or energy consumption, and (4) be immune to the hidden node problem. Though there have been numerous proposed MAC protocols for the IEEE 802.11 WLAN, few of them possesses all of the above four properties. In this paper, we propose a new MAC scheme satisfying all of above four properties. Our protocol can support real-time traffic and satisfy QoS requirements, and can achieve fairness among non-real-time nodes. Also, without using any centralized control, it can be easily deployed in MANET. An analytic model of the protocol's throughput has also been developed. We compare the protocol's throughput obtained from its analytic model and simulation to validate each other.

Keywords: IEEE 802.11, WLAN, MAC, QoS, Ad Hoc network

[1] **本文已在**IEEE International Conference on Communications (ICC) 2003 **刊登，詳細內容如附件。**

NICE - A Decentralized Medium Access Control Using Neighborhood Information Classification and Estimation for Multimedia Applications in Ad Hoc 802.11 Wireless LANs

Anderson Chen, Li-Chun Wang, Y.T. Su, Y. X. Zheng, and Bill Yang,
Dept. of Comm. Eng.
National Chiao Tung Univ.
Hsin-Chu, Taiwan 30050

David S.L. Wei
Dept. of CIS
Fordham Univ.
Bronx, NY 10458, USA

K. Naik
Dept. of E&CE
Univ. of Waterloo
Ontario, N2L 3G1, Canada

Abstract—**The desired properties of a medium access control (MAC) protocol in mobile ad hoc network (MANET) include (1) meet quality of service (QoS) requirements for real-time nodes, (2) be decentralized, (3) achieve fairness from viewpoint of throughput or energy consumption, and (4) be immune to the hidden node problem. Though there have been numerous proposed MAC protocols for the IEEE 802.11 WLAN, few of them possesses all of the above four properties. In this paper, we propose a new MAC scheme satisfying all of above four properties. Our protocol can support real-time traffic and satisfy QoS requirements, and can achieve fairness among non-real-time nodes. Also, without using any centralized control, it can be easily deployed in MANET. An analytic model of the protocol's throughput has also been developed. We compare the protocol's throughput obtained from its analytic model and simulation to validate each other.**

Keywords-IEEE 802.11, WLAN, MAC, QoS, Ad Hoc network

I. INTRODUCTION

IEEE 802.11-based wireless LANs have been increasingly in demand and will soon be prevalent. The desired properties of an IEEE 802.1l-based MAC protocol include: (1) meet QoS requirements for real-time nodes, (2) be decentralized, (3) achieve fairness from viewpoint of throughput or energy consumption, and (4) be immune to the hidden node problem. Though there have been several proposed MAC protocols for WLAN, to our knowledge, few protocols satisfie all of the above four properties.

In the literature, several QoS-guaranteed WLAN MAC protocols were proposed [1, 2, 3, 4, 5, 6, 7]. However, most protocols [1, 2, 3] designate one node as a central controller to coordinate the process of real-time data flows of the nodes within the radio broadcast range. The computational complexity of the bandwidth reservation process is very high and the battery power in the central controller may be consumed very soon. Besides, most protocols ignore fairness among the non-real-time nodes while meeting QoS requirements. There are also several protocols proposed with emphasis on high channel throughput and guaranteed fairness (see e.g. [8, 9, 10]). Though they are decentralized and some of them also consider the hidden node problem, they don't provide QoS guarantees, and neither was designed as an IEEE 802.11-compliant one. Also, the hidden node problem has been one of the main factors causing collision, thereby degradeing channel throughput. The standard IEEE 802.11 MAC protocol uses control packets, namely RTS and CTS, to avoid the hidden node problem. However, the MACprotocol to enhance the QoS requirement of 802.11 WLAN is still under development.

In this paper, we propose a new MAC scheme, namely neighborhood information classification and estimation (NICE) protocol, with all of the four desired properties. Based on a combination of the reservation and contention-based schemes, the proposed MAC scheme not only guarantees QoS requirements, but also provides the flexibility on bandwidth allocation between different traffic types. By overhearing the information exchanges, all the nodes within the same radio broadcast range are able to know the status of real-time traffic reservation and the number of neighboring nodes, thereby saving a lot of work in prediction and fully utilizing the channel capacity. The proposed NICE protocol also monitors the length of non-real-time traffic queue to give higher priority when the packet queue length of non-real-time traffic is larger than a predetermined threshold. During the operation of the proposed MAC scheme, our protocol does not rely on any central control mechanism. Thus, it can be employed in an ad hoc environment. The NICE protocol does not require each node to broadcast the beacon, while provideing the fairness from the viewpoint of energy consumption.

In Section II, we detail our proposed NICE protocol. An analytic model of the protocol's throughput is developed in Section III. In Section IV, we simulate the protocol using the ns2 simulator to study its performance in terms of channel throughput and access delay. Conclusions are given in Section V.

II. PROPOSED PROTOCOL

In this section, we detail our proposed NICE protocol. This protocol employs the IEEE 802.11 standard [11] as a subroutine for channel contention, but with a small modification in order to reduce the collision rate). We group data packets into two types, namely non-real time packets (nrt-packets), and real-time flows (rt-packets). For simplicity, a node that sends nrt-packets is called a nrt-node. Also, a node that sends rt-packets is called a rt-node. The MAC header field is the same as the 802.11 standard, except for rt-packets and the corresponding ACKs (rt-ACKs). The modified MAC header field for rt-packets and rt-ACKs are shown in Fig. 1.

The proposed protocol can be divided into four main procedures: observation, contention, reservation procedures and frame synchronization mechanism.

This work was jointly supported by the National Science Council, Taiwan under the contract 89-E-FA06-2-4, 90-2213-E-009-068, 91-2219-E-009-016 and EX-91-E-FA06-4-4.

0-7803-7802-4/03/$17.00 © 2003 IEEE

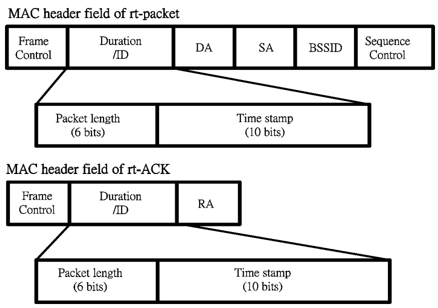

Fig. 1 MAC hearder field of real-time packet and it ACK in NICE protocol.

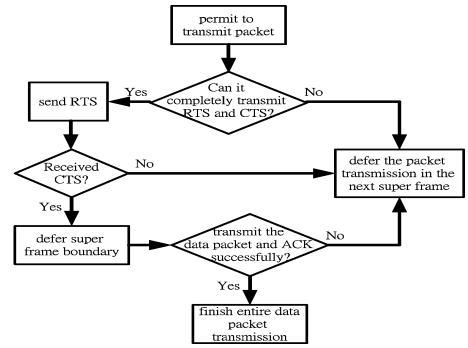

Fig. 2 Flow chart of FSM.

A. Observation Procedure

A node that just turned on its power or just moved in a new radio broadcast range has to observe the channel for a period longer than RP_{max} in order to know the system status of the radio broadcast range, where RP_{max} is defined as the maximum repetition period of a super frame in the PCF mode of the IEEE 802.11 standard. During this observation period, the node has to record what is overheard into a reservation table (RST) or a neighborhood information table (NIT). The RST records the reservation information carried in the MAC header field of rt-packets or rt-ACKs of different flows transmitting in current super-frame. The NIT records the time instance (D_TS) and source address of all the nrt-packets appearing in the current super-frame. This table provides the information about the number of active nrt-nodes within the region, which will be used to reduce the collision rate and improve the channel throughput in our design.

B. Contention Procedure

Our contention procedure consists of two phases, which is an extended version of 802.11 and 802.11e [12] standards. The first phase is identical to the DCF mode of 802.11 standard except that in our design nodes adopt different size of contention window (CW) for different types of packet. The nrt-nodes and rt-nodes use the values of (nrt-CW_{min}, nrt-CW_{max}) and (rt-CW_{min}, rt-CW_{max}) to control and limit the CW, respectively. In general, the priority of rt-packets is assigned to be higher than that of nrt-packet. Thus, we have the following relations for different CW boundaries:

$$\textbf{nrt-CW}_{\textbf{min}} > \text{rt-CW}_{\text{min}} \quad (1a)$$

$$\textbf{nrt-CW}_{\textbf{max}} \geq \text{rt-CW}_{\text{max}}. \quad (1b)$$

When a packet collides before, the corresponding rt-CW for rt-packet or nrt-CW for nrt-packet has the following relations:

$$\textbf{rt-CW} = \textbf{min(rt-CW}_{\textbf{max}}\textbf{, rt-CW}_{\textbf{min}} \times \textbf{(Num\_att-1))} \quad (2a)$$

$$\textbf{nrt-CW} = \textbf{min(nrt-CW}_{\textbf{max}}\textbf{, nrt-CW}_{\textbf{min}} \times \textbf{2}^{\textbf{(Num\_att-1)}}\textbf{)}, \quad (2b)$$

where **Num_att** is the number of attempts that the packet has tried. Once the transmission is successful, both rt-CW and nrt-CW are reset to their minimum value.

In the second phase, this packet gating mechanism decides whether the packet should be transmitted right away or deferred. First, the node needs to calculate the probability of transmission (*P_T*), based on the information obtained from the NIT table and gathered during the observation procedure. If the packet is permitted after this mechanism, it will get a green light for transmission. If the packet is deferred, the packet needs to restart the contention procedure again, and the contention procedure takes this packet as a collided one. The purpose of the gating mechanism is to reduce the collision rate based on the status of channel usage.

C. Reservation Procedure

Only nodes with rt-packets have to execute the reservation procedure. The node waits for a PCF inter-frame spacing (PIFS) period until the reserved time stamp (R_TS) in RST arrives. The R_TS indicates that the node has to transmit the rt-packet at the R_TSth time slot in the current super-frame. The value of R_TS is assigned in the previous rt-packet. Two rt-packets in the same super-frame are spaced with a PIFS period. The other information in RST is R_LEN, which is also in the unit of slot time. The value of R_LEN_i depends on the required bandwidth (ABR_i, average bit rate) of flow *i*.

D. Frame Synchronization Mechanism

Frame synchronization is another important issue in our proposed protocol. In the IEEE 802.11 or 802.11e standards, a point coordinator broadcasts a "beacon" signal as the start of a super-frame. However, from the viewpoint of energy consumption, broadcasting beacon signal wastes battery energy, especially for energy-sensitive environment such as MANET. Therefore, we propose a decentralized frame synchronization mechanism (FSM).

The proposed FSM filters the timing of the outgoing of nrt-packets in the nrt-nodes. Figure 2 shows the flow-chart representation of the proposed FSM. An nrt-node has to ensure that it can complete the RTS and CTS transmission before the end of a super-frame. Otherwise, this nrt-node needs to stop transmission in current super-frame and defer it to the next super-frame. Once the RTS and CTS packets are transmitted and received successfully, all the nodes in the region extend their super-frame boundaries to the time tick. After all the nodes extend their boundaries, they will not change the super-frame boundary any more. Even though the data packet has an error or there is no ACK, all the nodes still retain the extended super-frame boundary.

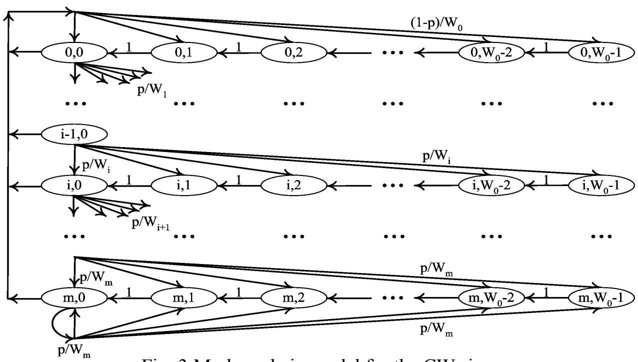

Fig. 3 Markov chain model for the CW size.

III. ANALYSIS

In this section, we analyze the throughput of our proposed protocol with mixed real-time and non-real-time traffic. We assume that (1) the channel is ideal, i.e. there is no packet error, (2) a fixed number of nrt-nodes are present in the region of radio broadcast scope, and each active node is always transmitting a packet, and (3) the "on/off" model is adopted for real-time traffic, where the periods of "on" and "off" state are exponentially distributed.

A. Previous case: Non-real-time traffic

Throughput analysis of non-real-time traffic is summarized as follows. First, define the normalized throughput, from [13], S, as the ratio of successful packet transmission time to the total transmission time. That is

$$\begin{aligned} S &= \frac{E[\text{payload transmitted in a slot time}]}{E[\text{length of a slot time}]} \\ &= \frac{P_s P_{tr} E[P]}{(1-P_{tr})\sigma + P_{tr}P_sT_s + P_{tr}(1-P_s)T_c}, \end{aligned} \tag{3}$$

where P_{tr} is the probability of at least one packet being transmitted in one slot, P_s is the probability of exactly one packet being transmitted in one slot, $E[P]$ is the average packet payload size, σ is the duration of an idle slot, T_s is the average successful transmission time, and T_c is the average collision time. Assume that there are n nodes contending for transmission, and each node transmits a packet in a slot with the probability τ. Then,

$$P_{tr} = 1-(1-\tau)^n \tag{4}$$

,and

$$P_s = \frac{n\tau(1-\tau)^{n-1}}{P_{tr}} = \frac{n\tau(1-\tau)^{n-1}}{1-(1-\tau)^n}. \tag{5}$$

Now, to find the probability that a node transmits a packet in a slot, denoted as τ, we consider a two-dimension discrete-time Markov chain shown in Figure 3. Denote $b(t)$ and $s(t)$ be the backoff time counter and the backoff stage at time t, respectively. Each state in the model represents the CW size in different backoff stage. Let the stationary probability of each state $b_{i,k} = \lim_{t\to\infty} P\{s(t)=i, b(t)=k\}$, where $i \in (0,m), k \in (0, W_i - 1)$

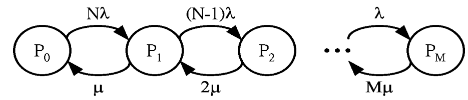

Fig. 4 Markov chain model for one type of real-time traffic.

Note that m is the maximum backoff stage and $W_i=2^iCW_{min}$. through conservation law in Markov chain, we can derive

$$b_{i,k} = \frac{W_i - k}{W_i} \cdot \begin{cases} (1-p)\sum_{j=0}^{m} b_{j,0} & i=0 \\ p \cdot b_{i-1,0} & 0<i<m \\ p\cdot(b_{m-1,0}+b_{m,0}) & i=m \end{cases}, \tag{6}$$

where p is the collision probability conditioned on a packet being transmitted. Following the approach in [13], $b_{0,0}$ can be expressed by

$$b_{0,0} = \frac{2(1-2p)(1-p)}{(1-2p)(W+1)+pW(1-(2p)^m)}. \tag{7}$$

Thus, we can express the probability τ as

$$\begin{aligned} \tau &= \sum_{i=0}^{m} b_{i,0} = \frac{b_{0,0}}{1-p} \\ &= \frac{2(1-p)}{(1-2p)(W+1)+pW(1-(2p)^m)}. \end{aligned} \tag{8}$$

From (8), one can observe that τ depends on the unknown conditional collision probability p. Note that $p = 1-(1-\tau)^{n-1}$. Thus, τ and p, need to be solved recursively by numerical methods.

B. Mixed real-time and non-real-time traffic

The analytic model presented in section III.A is for non-real-time traffic alone. In this part, we consider a mixed traffic model with both real-time traffic and data traffic. Rt-nodes and nrt-nodes generate real-time traffic and non-real-time traffic, respectively. The maximum number of rt-nodes and nrt-nodes is M and K. Let $p(t)$ represent the number of rt-nodes that request to transmit packets at time t. We assume that during a super-frame only one node joins or leaves the network, and real-time traffic and non-real-time traffic are generated independently.

We model the process $\{p(t)\}$ by a Markov chain as shown in Figure 4. Each state in the figure stands for the number of acting rt-nodes, i.e. which have packets waiting to be transmitted. The state probability can be expressed as $P_i = P\{p(t)=i\} = \binom{\mathcal{M}}{i}\rho^i P_0$ where ρ is the ratio of incoming probability and leaving probability, i.e. $\rho = \lambda/\mu$. Here, λ and ρ are the arrival rate and departure rate, respectively.

Denote $T(M, K)$ as the throughput of M rt-nodes and K nrt-nodes. Then, the sum of the throughput of real-time traffic and that of non-real-time traffic is given as

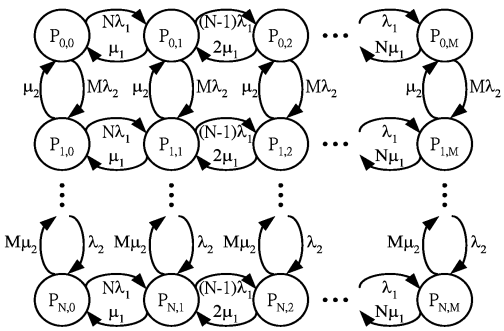

Fig. 5 Markov chain model for two types of real-time traffic.

$$
\begin{aligned}
T(M,K) &= E\begin{bmatrix}\text{throughput of } \mathcal{M} \text{ rt - nodes} \\ \text{and throughput of } \mathcal{K} \text{ nrt - nodes}\end{bmatrix} \\
&= \sum_{i=0}^{M-1} P_i \left\{ \frac{iL_{rt}}{L} T_{rt} + \frac{L - iL_{rt}}{L} T_{nrt}(K) \right\} \\
&= T_{nrt}(K) + \left(T_{rt} - T_{nrt}(K)\right) \sum_{i=0}^{M-1} P_i \frac{iL_{rt}}{L} \\
&= T_{nrt}(K) + \left(T_{rt} - T_{nrt}(K)\right) \frac{L_{rt}}{L} \frac{M\rho}{1+\rho},
\end{aligned} \tag{9}
$$

where L is the time span of a super-frame, and L_{rt} is the total time spent for transmitting an rt-packet in a super-frame. Note that L_{rt} includes a PIFS, a SIFS, an rt-packet, and an ACK. The period T_{rt} contains only one rt-packet. $T_{nrt}(K)$ represents the total amount of time spent for transmitting nrt-packets in K nrt-nodes.

Equation (9) can be extended to the case with a mixed traffic model containing multiple real-time traffic types, e.g. CBR and VBR. Under such a condition we then utilize a multi-dimension Markov chain. For example, consider two types of real-time traffic and one data traffic. Assume that N rt-nodes transmit one type of real-time traffic, M rt-nodes transmit the other type of real-time traffic, and K nrt-nodes transmit data traffic. Figure 5 shows the two-dimension Markov chain of this mixed traffic model. The state probability, $P_{i,j}$, can be expressed as $P_{i,j} = \binom{\mathcal{N}}{i}\binom{\mathcal{M}}{j}\rho_1\rho_2 P_{0,0}$, where $P_{0,0} = \frac{1}{(1+\rho_1)^{\mathcal{N}}(1+\rho_2)^{\mathcal{M}}}$. Note that ρ_1 and ρ_2 are the ratio of the probability of incoming probability and leaving probability for type 1 real-time traffic and type 2 real-time traffic, respectively. The total throughput, $T(M, N, K)$, can be rewritten as

$$
\begin{aligned}
T(M,N,K) = T_{nrt(K)} + (T_{rt1} - T_{nrt}(K)) \frac{L_1}{L} \frac{M\rho_1}{1+\rho_1} + \\
(T_{rt2} - T_{nrt}(K)) \frac{L_2}{L} \frac{N\rho_2}{1+\rho_2},
\end{aligned} \tag{10}
$$

where all the parameters have been defined in (9).

IV. PERFORMANCE SIMULATION

In this section, we discuss the performance of our protocol in terms of channel throughput and access delay via simulations using the ns-2 simulator. The simulation parameters are listed in Table 1.

Table 1:

Parameter Names	Parameter Values
Cover Range	10m×10m
Data Rate	2 Mbps
Simulation time	100 sec
Unit slot time	20 μsec
SIFS/PIFS/DIFS	10/30/50 μsec
Min./Max. CW for nrt-packets	31/1023
CW for out of queue threshold	15
Min./Max. CW for rt-packets	7/31
Queue length threshold	20
Node expire threshold	100 super-frames

A. *Traffic Model*

We consider three types of traffic models in the simulation:

- *Voice traffic model*: The voice traffic is modeled as a Poisson process with "on" and "off" states. In the "on" state, the simulator continuously generates 164-byte packets every 20.48 ms, i.e. 64 Kbps. In the "off" state, the simulator stops creating packets. The duration of the "on" state and "off" state follows the exponential distribution. The average duration of the "on" and "off" state are 1 and 1.3 seconds respectively.
- *FTP data traffic model*: The FTP traffic model continuously generates constant sized packets. That is, after one node finishes transmitting a packet, it immediately attempts to transmit another packet.
- *Telnet data traffic*: In the simulation, the generation of a Telnet packet follows Poisson process. For simplicity, the packet size is assumed to be constant for all scenarios, but the arrival rates for different scenarios are different.

B. *Performance Measurements*

Our protocol is evaluated in terms of throughput, fairness, and packet dropping rate.

- *Normalized throughput*: The normalized throughput is defined as the ratio of the number of successfully received useful data bits to the total number of transmitted bits.
- *Fairness*: In evaluating the fairness, we collect the delay of all the packets in all the connections, and calculate the standard deviation of all the connections. We then illustrate the results of the connection that has the maximum value of standard deviation among all the connections.
- *Dropping rate*: The dropping rate is defined as the ratio of the number of *dropped* rt-packets to the total number of transmitted rt-packets. If an rt-packet is received longer than 2 RP_{max}, i.e. 40.96 msec, the packet is considered to be a dropped packet.

C. Numerical Results

Figure 6 shows the normalized throughput with mixed data and voice traffic. We consider the case when 10 nodes send voice traffic to 10 other corresponding nodes, and the other 40 nodes send FTP data packets. From the figure, one can see that our analysis, (7), and simulation results are very close especially for small packets. Even for a large packet size, the discrepancy between analysis and simulation is still less than 3%.

Figure 7 compares the fairness of the NICE protocol with the IEEE 802.11 standard with mixed Telnet data and voice traffic. Figure 8 shows the performance comparison of NICE and the IEEE 802.11 standard in terms of dropping rate with voice traffic. As shown in the figure, the dropping rate of NICE is almost negligible, while the dropping rate of the IEEE 802.11 standard is higher than 40% at heavy traffic load. For the IEEE 802.11 standard, every voice packet needs to contend for a time slot with other nodes. However, with the help of the reservation procedure in our protocol, each rt-node can reserve time slots for the data flow of voice packets.

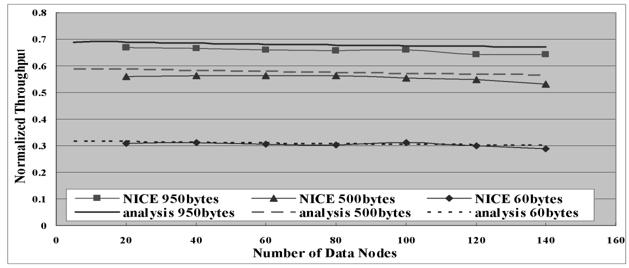

Fig. 6 Throughput comparison of the proposed NICE protocol with mixed FTP data traffic and voice traffic.

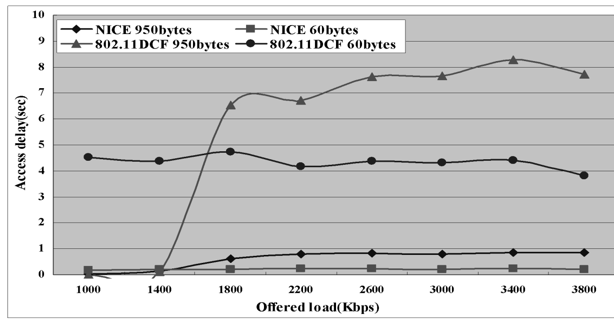

Fig. 7 Maximum standard deviation of access delay with mixed Telnet data and voice traffic.

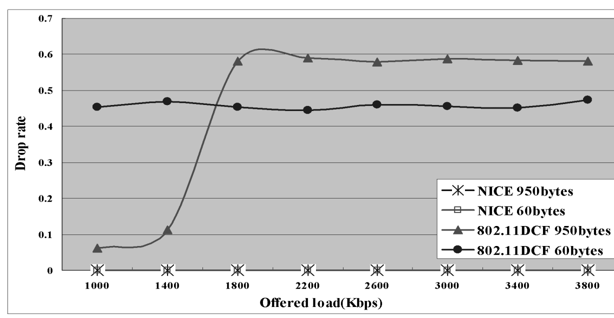

Fig. 8 Dropping rate of voice traffic in a mixed Telnet data and voice traffic.

V. CONCLUSIONS

We have proposed a new MAC protocol, called Neighborhood Information Classification and Estimation (NICE), for multimedia applications in MANET. Without using a central control, the NICE MAC protocol outperforms the standard IEEE 802.11 DCF mode in terms of throughput, mean access delay, fairness, and the QoS guarantee for the mixed data and voice traffic. In the NICE protocol, we have proposed four new mechanisms: (1) dynamic gating mechanism for throughput enhancement; (2) stall avoidance mechanism for achieving fairness; (3) linear backoff scheme for reducing the dropping rate of real-time traffic; and (4) reservation procedure for QoS guarantee. Via analysis and simulation by ns-2 simulator, we demonstrate that the throughput performance of NICE is at least 50% better than the IEEE 802.11 standard. The mean value and maximum standard deviation of access delay are 8 times less. The dropping rate of voice traffic of NICE is almost negligible, but the dropping rate of the IEEE 802.11 standard is at least 40%.

REFERENCES

[1] Shiann-Tsong Sheu and Tzu-Fang Sheu, "DBASE: A Distributed Bandwidth Allocation/Sharing/Extension Protocol for Multimedia over IEEE 802.11 Ad Hoc Wireless LAN," IEEE INFOCOM 2001.

[2] Shiann-Tsong Sheu, Tzu-Fang Sheu, Chih-Chiang Wu, and Jiau-Yu Luo, "Design and Implementation of a Reservation-base MAC Protocol for Voice/Data over IEEE 802.11 Ad-hoc Wireless Networks,"IEEE ICC 2001.

[3] Malathi Veeraraghavan,Nabeel Cocker, and Tim Moors,"Support of Voice Services in IEEE 802.11 Wireless Lans," IEEE INFOCOM 2001.

[4] J. L. Sobrinho and A. S. Krishnakumar, "Quality-of-Service in Ad Hoc Carrier Sense Multiple Access Networks," IEEE Journal on Selected Areas in Communications, 17(8):1353-1368, Aug. 1999.

[5] A. Veres, A. T. Campbell, M. Barry, and L. Sun, "Supporting Service Differentiation in Wireless Packet Networks Using Distributed Control," IEEE Journal on Selected Area in Communication, 19(10):2081-2093, Oct. 2001.

[6] G. Anastasi and L. Lenzini, "QoS Provided by the IEEE 802.11 Wireless LAN to Advanced Data Applications: a Simulation Analysis," ACM Wireless Networks, Vol. 6, pp. 99-108, June 2000.

[7] A. Lindgren, A. Almquist, and Olov Schelen, "Evaluation of Quality of Service Scheme for IEEE 802.11 Wireless LANs," in Proceedings of the 26th Annual IEEE Conference on Local Computer Networks (LCN 2001), Nov. 2001.

[8] T. Nandagopal, T. E. Kim, X. Gao, and V. Bharghavan, "Achieving MAC Layer Fairness in Wireless Packet Networks," MobiCom 2000, pp. 87-98.

[9] N. H. Vaidya, P. Bahl, and S. Gupta, "Distributed Fair Scheduling in a Wireless LAN," MobiCom 2000, pp. 167-178.

[10] L. Bononi, M. Conti, and L. Donatiello, "Design and Performance Evaluation of a Distributed Contention Control (DCC) Mechanism for IEEE 802.11 Wireless Local Area Networks," WOWMOM 98, pp. 59-67.

[11] IEEE Std. 802.11, Part II: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 1999 Edition.

[12] IEEE Std. 802.11e/D3.2, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Medium Access Control (MAC) Enhancement for Quality of Service (QoS), Aug. 2002

[13] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," IEEE Journal on Selected Areas in Communications, Vol. 18, no. 3, pp. 535-547, Mar. 2000.