May 2005

End-to-end Application Performance Impact on Scheduler in CDMA-1XRTT Wireless System

Bong Ho Kim
University of Pennsylvania

Insup Lee
University of Pennsylvania, lee@cis.upenn.edu

Kelvin Chu
University of Pennsylvania

End-to-end Application Performance Impact on Scheduler in CDMA-1XRTT Wireless System

Abstract
User-perceived application-level performance is very important to the adoption and success of 3G wireless services and infrastructure. This paper illustrates the end-to-end application performance when connecting through a CDMA-1XRTT network and compares the performance to that of a dialup connection. The results show that the application performance in CDMA-1XRTT may depend on the application traffic behavior. While many applications perform quite well in CDMA-1XRTT, others such as Microsoft Outlook perform below expectation and may even be slower than a dialup connection. This paper proposes an enhanced application sensitive scheduling algorithm to improve the performance of such applications. The enhanced algorithm also provides additional benefits for applications with similar behavior as Outlook. These include additional channel code availability and channel switching process reduction by as much as 60%.

Keywords
CDMA, Scheduler, Performance

Comments

End-to-end Application Performance Impact on Scheduler in CDMA-1XRTT Wireless System

Bong Ho Kim, Insup Lee, Kelvin Chu
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, U.S.A
{kimbong, lee}@cis.upenn.edu

***Abstract*— User-perceived application-level performance is very important to the adoption and success of 3G wireless services and infrastructure. This paper illustrates the end-to-end application performance when connecting through a CDMA-1XRTT network and compares the performance to that of a dialup connection. The results show that the application performance in CDMA-1XRTT may depend on the application traffic behavior. While many applications perform quite well in CDMA-1XRTT, others such as Microsoft Outlook perform below expectation and may even be slower than a dialup connection. This paper proposes an enhanced application sensitive scheduling algorithm to improve the performance of such applications. The enhanced algorithm also provides additional benefits for applications with similar behavior as Outlook. These include additional channel code availability and channel switching process reduction by as much as 60%.**

Keywords-component; CDMA, Scheduler, Performance

I. Introduction

3G wireless networks are expected to offer high data rates and data services for both individual user and the network. Two main radio transmission technologies that meet the desired data rates and other requirements are (1) CDMA2000 by the TIA of USA and, (2) UMTS (W-CDMA) by the ETSI of Europe. 3G wireless networks are currently deployed in the United States and Korea using the CDMA-1XRTT technology and data-only enhancements to CDMA-1XRTT have been standardized in the CDMA-EVDO standard with speeds up to 2Mbps [1]

This paper uses the CDMA-1XRTT as the backdrop of the analysis of the user achievable practical data rate in the delivery of wireless Internet in 3G wireless networks. As 3G networks provide pervasive Internet access, good performance of the wireless links will be critical for end user satisfaction. In order to mitigate the effects of losses, 3G wireless systems use extensive local retransmission mechanisms, e.g. link layer retransmission protocols such as RLP and RLC are used in CDMA-1XRTT [2] and UMTS [3]. In addition, intelligent channel state based scheduling has been introduced to increase throughput. The range of data rates over the air is from 9.6kbps up to 153.6kbps, which is realized by dynamically assigning an additional traffic channel (Supplemental Channel – SCH) to the packet data call. This new channel is only assigned when the data activity requires more bandwidth than is provided by the raw data rate dedicated channel. This high data rate channel (SCH) can vary both in duration of assignment and channel rate [4]. The scheduling of the SCH is a critical component of the air interface capacity of a CDMA-1XRTT system. It is the responsibility of the BSC to allocate and de-allocate the SCH efficiently so that SCH idle time is minimized and the usage of the air interface is maximized. This translates into more revenue for the operator, as there is a direct correlation between revenue and the usage of the air interface. The reference [5], [6] and [7] are only a few of researches to increase the utilization of the air interface.

A high speed data burst assignment consists of: (1) physical channel code resources that the mobile station shall decode during the data burst, (2) the current amount of data in the buffer and the rate at which data is arriving, (3) the data transmission rate employed on those physical code channels; and (4) other RF parameters. One critical information that cannot be ignored is the application protocol behavior because the scheduler may allow the SCH channel to expire even though a burst of packets is coming soon. A scheduler using only backlog buffer size may not provide fair channel allocation for different applications because of the variability in the application protocol behaviors. Different protocols generate different patterns of packet flow. Since such a scheduling algorithm does not distinguish between different applications, the application, which is not properly matched with the scheduler, will experience lower performance than those that properly match. A good scheduling algorithm must provide fairness among users and increase system throughput, but it may not provide fairness among applications. This will directly impact the end-to-end application performance that end-users mostly care about.

This paper compares some of typical CDMA-1XRTT application performance measurement results (Web, FTP, and Outlook Email) from a user's point of view to Dialup networks. To understand each application behavior and identify potential problems, a CDMA-1XRTT simulator is modeled and a network emulator is set up with actual application severs and clients. This paper also presents a detailed discussion of the application behavior and identifies where the unfairness can come from. Finally, this paper compares the application performance of a generic scheduling algorithm to the proposed application-sensitive scheduling algorithm. This paper does

0-7803-8887-9/05/$20.00 (c)2005 IEEE

not intend to provide an optimized algorithm as it is beyond the scope of this paper. The simple scheduling algorithm used is generic and does not represent any specific manufacturer's scheduling algorithm.

II. THE PROBLEM

In order to compare the application performance in CDMA-1XRTT and a dialup service, we used a commercial WSP network and a dialup ISP. Since both of the measurements are taken in an uncontrolled environment and the measurement results can vary with different configurations and environment, the measurement results or the comparison in this paper do not directly compare these technologies and should not be interpreted as an overall performance of each technology.

Figure 1 (c) shows the average ping delay with various packet sizes from 32 bytes to 1300 bytes. Ping delays with CDMA-1XRTT connection increases relatively linearly when the ping packet size is increased. Ping delays with dialup connection, however, are rather constant with larger ping packet size. This is due to the fact that (1) the radio frame error rate and RLP retransmission that does not exist in dialup connection and, (2) the lower channel rate assigned by the scheduler. We also measured Web, Email and FTP application throughput performance and the results are shown in Figure 1. For the Web page access performance test, the Yahoo home page is used and for the email performance, Outlook client and MS Exchange server are used. The results indicate that the CDMA-1XRTT system shows a throughput of 81 kbps on the average for FTP, and 53 kbps throughput for the Web page access which are 3.3 and 2.5 times respectively higher than the Dialup connection. Email application performance results, however, is about 1.5 times higher than the Dialup connection.

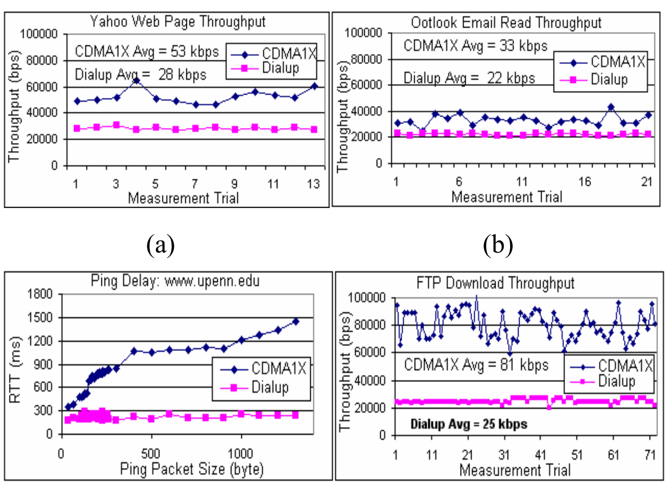

Figure 1 CDMX-1XRTT and dialup measurement

This improvement in throughput for the FTP and Web browsing would be attractive to the dialup users. However, the performance of some applications, such as Outlook, is comparable to the Dialup network. The overall application performance can be improved by optimizing some of configurable applications and network parameters. The bigger issue identified from the tests is that the Web browsing performance improvement is not as good as that of the FTP and the Outlook email performance improvement is even worse. This seems that certain applications have limited access to the higher channel rate even if the radio resource is available. One of the critical issues resulting in unequal application throughput is because applications generally are not designed to interact with a particular protocol layer.

III. THE PROPOSED APPROACH

A. *CDMA2000 Wireless Network Simulation Model*

Figure 2 shows a user-plane simulation model of an end-to-end reference connection through a CDMA-1XRTT wireless network. A base station is connected to a Base Station Controller (BSC), which performs specific functions such as soft handoff, power control, radio resource assignment, etc. It also performs radio link layer retransmission using the RLP protocol. In the CDMA-1XRTT system, the BSC is connected to a PDSN using a GRE tunnel and the PDSN terminates the PPP connection from the user mobile device.

In this architecture, the BSC receives PPP/IP packet through the GRE tunnel from the PDSN. The BSC fragments the packet into a number of radio frames and then transmits them using the RLP protocol. Some of the radio frames will be retransmitted according to the RLP layer retransmission mechanism. The user device receives the radio frames and if it discovers lost of radio frames, it requests retransmission using RLP protocol. The RLP retransmission occurs within the wireless access area. The circuit network and voice traffic reference connection are not shown since this paper only considers data applications. The simulator models all protocol layers from the physical through the application layer and models details of the packet handling characteristics of each network element along the path. The simulation model predicts application-level performance metrics such as response time, packet loss, and jitter.

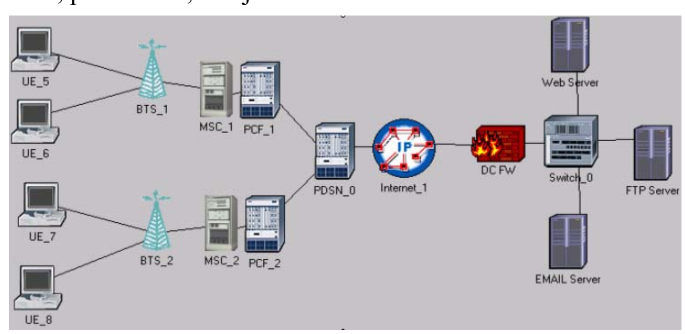

Figure 2 CDMA-1XRTT network simulator architecture

B. *Wireline Network Emulator Setup*

The CDMA-1XRTT wireless simulator described above can provide a very detailed application performance when the application with well-known behavior or protocol, such as FTP and HTTP, is modeled in the simulator in detail. Outlook and Exchange Server, however, use a Microsoft proprietary protocol (Messaging Application Programming Interface) in the application layer and the description is not available. To understand the fundamental behavior of the Outlook application and to accurately model it, a wireline based network emulator was used to capture a packet and release it to the network after a specified delay time. This is not a packet monitoring type of system, but rather it is a packet intercept system. This system provides a completely controllable application-testing environment and can also be used to emulate wireless network.

IV. APPLICATION BEHAVIOR ANALYSIS

Web browsing, FTP and Email applications are the most popular applications in wireline and will constitute a significant fraction of the packet data traffic in 3G wireless systems; Outlook and Exchange Servers are selected for the e-mail application. At the session layer, HTTP is used by web browsers while the MAPI protocol is used by both Outlook and Exchange Servers. Although there are significant differences in the characteristics of HTTP, FTP, and Email sessions, the underlying transport protocol used in HTTP, FTP, and MAPI is TCP. The detailed traffic model for HTTP and FTP applications can be found in [8] and Outlook behavior is described in this section based on packet analysis. The MAPI is a de facto Microsoft system standard for messaging and workflow applications. The MAPI protocol provides a feature-rich set of functionality for accessing Exchange. However, these benefits seem to come with performance degradation in CDMA-1XRTT systems.

A. Outlook Application Traffic Characteristics

First of all, Outlook depends upon Remote Procedure Call (RPC) for its transport and uses a somewhat chatty protocol. During the Outlook invoke phase, a lot of the chattiness is observed. There are eleven active TCP connections during the invoking phase. Each Outlook email transaction is consisting of multiple MAPI segment transactions in series and each MAPI fragment is segmented into smaller segments again. During the test, most of the time, the maximum MAPI fragment was 16896 bytes and this information is indicated in the first packet of a MAPI fragment. The maximum size of the transmit and receive MAPI sub-fragment is negotiated during the invoking phase (5840 bytes was observed). Outlook finishes all the ACK packet transmission for the current MAPI segment and the Exchange server waits for the MAPI fragment completion indication packet before sending the next one. The last packet in the MAPI fragment sets the "PUSH" bit in the TCP packet to transmit all of the packets in the TCP buffer to the application layer at the receiver side. In the actual CDMA-1XRTT system the average MAPI fragment gap was about 1.2 s, but gaps longer than 2 sec were sometimes observed.

The result clearly shows that the Outlook application behavior is completely different from the FTP behavior. The total packet silence gap during a single Email transaction contributes from 2% to 15% of the total response. The packet silence gap between the MAPI segments can be approximated as following.

Packet silence gap = (last MAPI packet size + server response packet size) * link rate + 2 * e2e network delay + server processing time + client processing time. (1)

In a wireline network with shorter RTT and higher link rate, the MAPI protocol does not impact application performance. Even in the wireless network with longer RTT and lower link rate, there may not be much performance degradation (without considering packet loss) since the effect of low link rate is more dominant than RTT and user application performance expectation would be low. However, the silence gaps can cause significant performance degradation in wireless systems, especially in the CDMA-1XRTT system, since CDMA-1XRTT networks allocates the channel rate based on link usage. The next section describes how the MAPI protocol can impact the application performance in CDMA-1XRTT system.

B. Interaction between Application and Channel Rate Assignment

This section describes how application behavior can interact with the channel rate assignment algorithm using the FTP and Outlook. The channel rate diagram is drawn from the fundamental channel rate assignment mechanism description. Figure 3 (a) matches the link rate assignment and the FTP packet trace. Since the FTP sends packets without interruption, the scheduler can assign a higher SCH channel rate one after another whenever resources are available. This scenario is shown in Figure 3 (a) with the continuous 153.6 kbps SCH channels. Figure 3 (c) shows the FTP throughput every 1 sec interval from a measurement from CDMA-1XRTT as an example, which shows a peak of 140 kbps and maintains higher channel rate. However, Figure 3 (b) shows that the Outlook application cannot get full advantage of high SCH channel rate. Since the MAPI protocol interrupts the packet flow between MAPI fragments, these interruptions prevent the higher data rate request by emptying the buffer in the RLP layer. Every MAPI fragment may start from the 9.6 kbps fundamental data rate and ramp up to a higher data rate while the same MAPI fragment is transmitted.

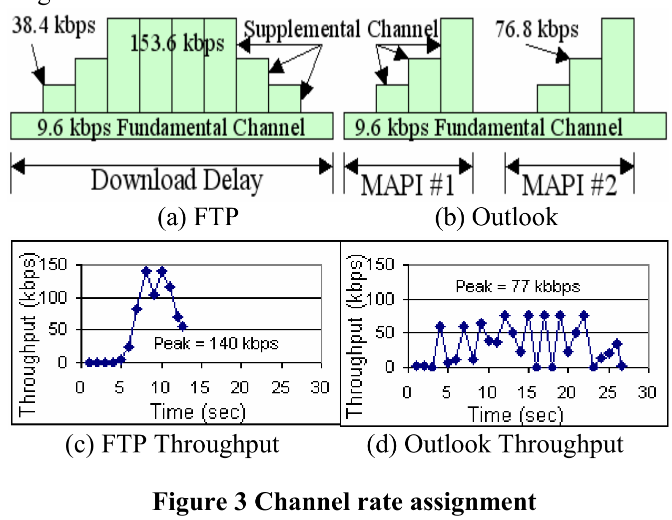

Figure 3 Channel rate assignment

This can be called a slow start channel rate assignment similar to the TCP slow start. This slow start channel rate assignment can exist even under no overload in the system with best air interface condition because this behavior comes from the application. Figure 3 (d) shows the Outlook throughput every 1 sec interval from a measurement from CDMA-1XRTT as an example, which shows a fluctuating throughput peak of 77 kbps.

V. CDMA-1XRTT SIMULATION

This section presents the CDMA-1XRTT simulation results for Web, FTP, and Email application using two different scheduling algorithms: a simple generic algorithm and an enhanced application-sensitive one.

A. Simulation Scenarios and Parameters

It is assumed that each of the simulation scenarios contains one application session at a time. The transmission time interval (TTI) for the data application is set to 20 ms and 1% ~ 5% of the frame error rate for the forward and reverse link based on the assigned SCH channel rate. The maximum allowable retransmission to recover the frame errors between the RLP layers is limited to three, and those packets not recovered by the RLP layer rely on the recovery mechanism in TCP. One of the key assumptions for the scheduler is to assume a scenario of a single user with no power and bandwidth (i.e. Walsh code) limitation. Thus the scheduler always grants the required channel rate. This is not quite practical since there would be resource contention and frequent wireless link condition changes but the assumption is sufficient for this study, because the focus is on the interaction between the scheduler and application behavior.

The Simulation is verified by comparing its application performance to that of a generic scheduling algorithm in terms of application layer throughput with the measurement through a WSP's CDMA-1XRTT network. Figure 4 shows that the throughput average and standard deviation for the Outlook, Web browsing and FTP application. As noted previously, FTP shows highest throughput and the Outlook application experiences the worst throughput. The average throughput difference between the measurement and the simulation is less than 5% and more importantly the scheduler in the simulator behaves very close to the real CDMA-1XRTT system.

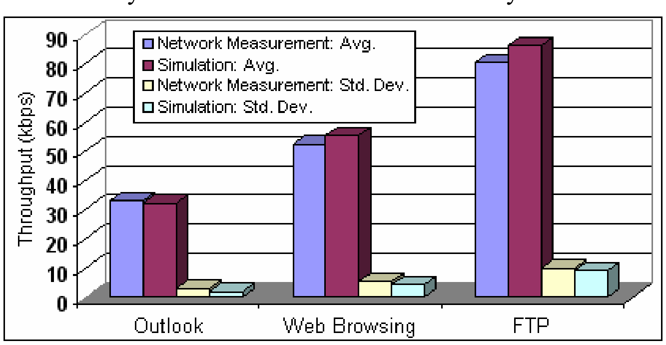

Figure 4 Application throughput comparison: measurement vs. simulation

B. Scheduling Algorithm

In the case of a simple generic scheduler, the required SCH channel rate and the duration are decided based on the backlog in the RLP layer. If the backlog is large, then a higher channel rate is assigned and the duration is calculated based on the backlog and the assigned channel rate. Once the buffer becomes empty and previous SCH channel is expired, the SCH is released until next SCH channel request comes.

However, the proposed application-sensitive scheduler, instead of dropping the SCH channel when the RLP buffer is empty, the SCH channel is maintained at a one step lower channel rate or at least at the lowest SCH channel rate for a certain duration (= Grace_Timer minus time passed after the last packet transmission). The Grace_Timer value can be determined on the fly by analyzing the packet inter-arrival time. In order to analyze the sensitivity of the scheduler, it is necessary to use a fixed Grace_Timer. If there is no packet activity within the Grace_Timer duration, then the SCH channel will be released; otherwise, the application session continues to use the SCH channel without the SCH channel re-establishment process. Even if a long Grace_Timer is used and there is no packet activity, the radio channel can still be released by the radio resource management process upon the expiration of an inactivity timer (if it is enabled).

C. Simulation Resuls

The proposed algorithm is sensitive to the behavior of applications, such as Outlook, and reduces their response time. Figure 5 shows that the response time for all of the application is improved with the proposed scheduling algorithm. The focus of this paper is to show that there is significant throughput improvement among some applications, such as Outlook, when using the enhanced algorithm.

Figure 5 shows that there is a 13-17% improvement in Outlook, 7-8% in Web browsing and 5-6% in FTP with the application sensitive scheduler. This indicates that the proposed scheduler treats each application differently based on the application behavior. It also shows that the improvement is leveling off after a certain length of a Grace_Timer, i.e., 2 sec for Outlook and 0.5 sec for Web browsing and FTP because it is not likely to have packet inter-arrival time longer than this duration. The remaining discussion will therefore focus only on the effect of the proposed algorithm on those applications such as Outlook.

However, the reduction in response time is achieved at the expense of additional bandwidth assignment. Figure 6 shows that the additional amount of bandwidth assigned to each type of application. To support Outlook, approximately 25% more bandwidth has been assigned and less than 4% for other applications. Since the network is usually not 100% utilized, the increase in bandwidth assignment may not affect the total number of users in the system. If the utilization becomes heavy, a policy can be established on how the additional bandwidth can be assigned.

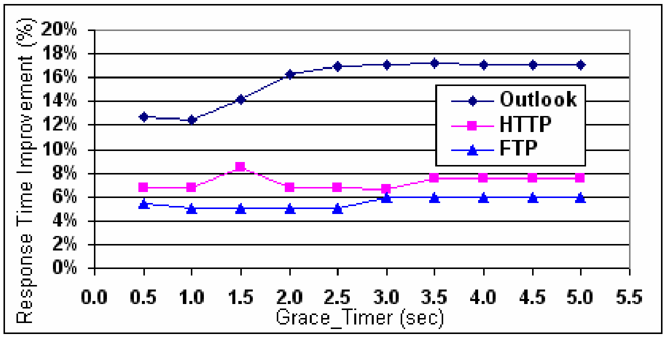

Figure 5 Application performance improvement with the application-sensitive scheduling algorithm

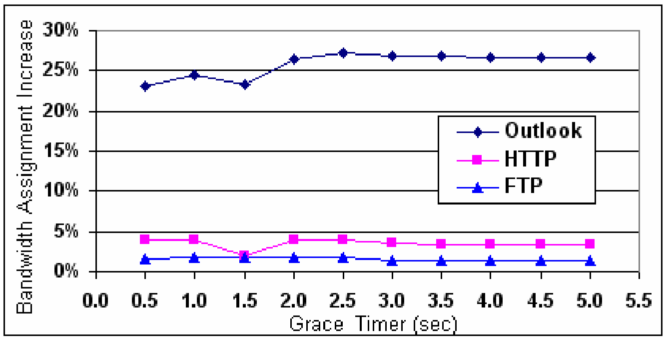

Figure 6 Amount of additional bandwidth assignment

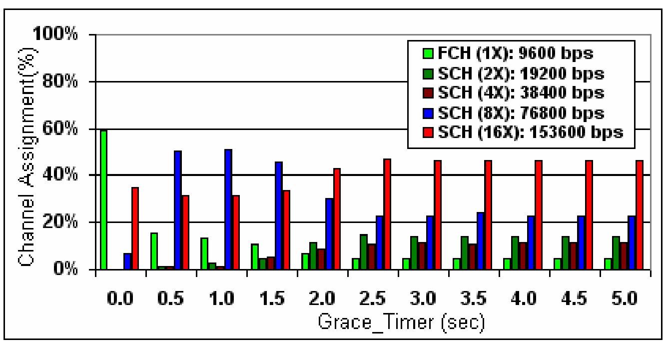

Figure 7 Distribution of channel rate assignment

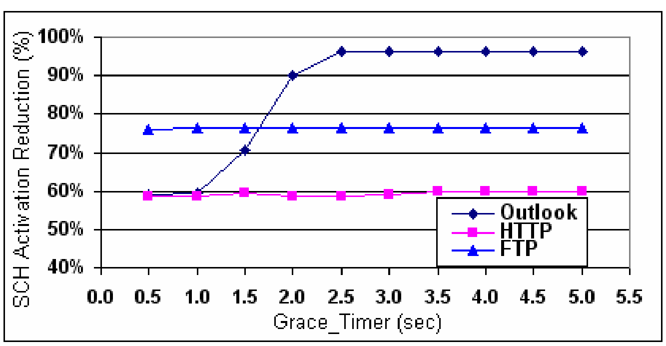

Figure 8 Channel switching process reduction

The proposed algorithm allows more channel codes to be available. Channel code is one of the limiting factors affecting the total number of simultaneous users in a system. The proposed scheduler tends to utilize more of the mid-band channel rates, such as 19.2kbps (2X), 38.4 kbps (4X) and 76.8 kbps (8X), than the generic scheduler and reduces the utilization of both the 9.6 kbps (1X) and 153.6 kbps (16X). Figure 7 also shows that the utilization of the 8X channel is much higher than that of the 16X channel when the Grace_Timer is <= 2 sec. As noted previously, there is very little likelihood that the inter-arrival time will be greater than 2 sec.

In addition to the benefits above, this scheduler also has an effect of reducing the system-processing load, which is due to the switching between FCH and SCH. Figure 8 shows that the proposed scheduler will reduce the channel switching processing by at least 60%.

VI. Conclusions and Recommendation

This paper illustrates that the generic scheduler is not application sensitive and therefore certain applications such as Outlook will not benefit from high speed network such as CDMA-1XRTTRTT. This paper proposes an algorithm which is sensitive to the behavior of the applications. This will reduce the response time of such applications. However, there is a tradeoff between the response time reduction and the total assigned bandwidth. The proposed algorithm also increases the number of the available channel codes as well as reducing the channel switch between FCH and SCH by at least 60%. The proposed scheduler presented in this paper recognizes the application behavior and is able to allocate channels to various applications fairly. It improves the performance of such applications as Outlook significantly compared to the generic scheduler, which does not recognize this important parameter. The enhanced scheduling algorithm should be optimized in order to be practical, and it will be presented in the future.

References

[1] 3GPP2, “cdma2000 High Rate Packet Data Air Interface Specification”, Dec 2000.

[2] TIA/EIA/IS-707-A-2.10, “Data Service Options for Spread Spectrum Systems: Radio Link Protocol Type 3”, January 2000.

[3] 3GPP, “RLC Protocol Specification (3G TS 25.322)”, 1999.

[4] 3GPP2 C.S0002-C, Physical Layer Standard for CDMA2000 Spread Spectrum Systems

[5] N.Joshi, S. Kadaba, S. Patel, and G. Sundaram, “Downlink Scheduling in CDMA Data Networks”, Proceedings of ACM Mobicom, August 2000.

[6] A. Jalali, R. Padovani, and R. Pankaj, “Data Throughput of CDMA-HDR, a High Efficiency Data Rate Personal Communication Wireless System”, Proceedings IEEE VTC, 2000.

[7] P. Bhagwat at al, “Enhancing Throughput over Wireless LANs Using Channel State Dependent Packet Scheduling,” in Proc. IEEE INFOCOM’96.

[8] 3GPP2-C50-EVAL 2001 022-0, HTTP and FTP Traffic Models for 1xEV-DV Simulations, Lucent, Motorola, Qualcomm, Samsung