This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the ICC 2008 proceedings.

Analysis of Uplink Traffic Characteristics and Impact on Performance in Mobile Data Networks

Mira Yun*, Yanxia Rong*, Yu Zhou*, Hyeong-Ah Choi*, Jae-Hoon Kim†, JungKyo Sohn‡ and Hyeong-In Choi‡
*Department of Computer Science, George Washington University, Washington, DC
†Access Network & Mobile Terminal R&D Center, SK Telecom, Seoul, Korea
‡Department of mathematics, Seoul National University, Seoul, Korea
Email: {mirayun, yxrong, yuzhou, hchoi}@gwu.edu, jayhoon@sktelecom.com, {jgsohn, hichoi}@snu.ac.kr

***Abstract*—User-generated content (UGC) also known as user-created content (UCC) refers to various kinds of media contents that are produced by end users. A recent trend in the use of the Internet exhibits that data traffic from UGC is rapidly growing with the potential to create a huge amount of uplink traffic for wireless operators. As the reality of UGC's scope and power is becoming crystalized, modeling and analysis of uplink traffic has just begun to receive growing attention in the wireless community. Consequently, we have also analyzed live uplink traffic traces obtained by monitoring 3G networks of a mobile data service provider (SK Telecom in Korea). Our statistical analysis shows the self-similarity in this traffic trace. In order to evaluate the impact of this characteristics on mobile data networks, we use the WiMAX module available in OPNET software. Our trace-driven simulation results indicate burstiness in the aggregated traffic received at a base station (BS) as well as in the traffic generated at each subscriber station (SS). The impact of the data burst at SSs is shown to be negligible when the overall traffic load is relatively small but become significant when a large volume of delay-sensitive traffic is generated in many SSs even if the overall network load is less than the network capacity.**

I. Introduction

User-generated content (UGC) also known as user-created content (UCC) refers to various kinds of media contents that are produced by end users. A recent trend in the use of the Internet exhibits that data traffic from UGC is rapidly growing with the potential to create a huge amount of uplink traffic for wireless operators [1]. As the reality of UGC's scope and power is becoming crystalized (e.g., YouTube, Facebook, Wikipedia, MySpace, and Flickr), modeling and analysis of uplink traffic has just begun to receive attention in the wireless research community [2]. We have also analyzed live uplink traffic traces obtained by monitoring 3G networks of a mobile data service provider (SK Telecom in Korea). A recent survey on the participative web and UCC [1] shows that about 96% of Korean Internet users know or have heard about UCC, and more than 50% of them have uploaded their UCCs to a popular social networking site Cyworld, operated by the SK Telecom.

Our statistical analysis shows that this trace exhibits a certain degree of self-similarity. In order to evaluate the impact of this traffic characteristics on 3G/4G mobile data networks, we use the Worldwide Interoperability for Microwave Access (WiMAX) module available in OPNET software [3]. Our trace-driven simulation results indicate burstiness in the aggregated traffic received at a base station (BS) as well as in the traffic generated at each subscriber station (SS). The impact of the data burst at SSs is shown to be negligible when the overall traffic load is relatively small but becomes significant when a large volume of delay-sensitive traffic is generated in many SSs even if the overall network load is less than the network capacity.

The paper is organized as follows. In the following, we present a brief review of mobile network evolution towards uplink enhancement. In Section III, we analyze live traffic traces of uplink traffic and discuss its self-similar characteristics. In Section IV, the impact of the data burst on the network performance is discussed based on trace-driven OPNET simulations using the WiMAX module. Conclusions are given in Section V.

II. Mobile Network Evolution: Towards Uplink Enhancement

Fast growth in cellular usage with emerging multimedia applications have led to the requirement for new 3G cellular telecommunication networks. To deal this requirement, two new partnership projects, 3rd Generation Partnership Project (3GPP) and 3GPP2, were established in 1998 [4][5]. 3GPP is developing 3G standard for Global System for Mobile (GSM) based system such as General Packet Radio Service (GPRS) and Universal Mobile Telecommunications System (UMTS)(or Wideband Code Division Multiple Access (WCDMA)) and 3GPP2 is focusing on Interim Standard (IS)-95 based CDMA system such as CDMA2000. As high data rate services such as video transmission and other data services became popular, both 3GPP and 3GPP2 introduced downlink enhancement technologies of each 3G system (WCDMA, CDMA2000), which are High Speed Downlink Packet Access (HSDPA)(up to 10Mbps) and CDMA 1x Evolution Data Only (EvDO) (up to 2.4 Mbps in Rev. 0) respectively [6][7].

Initially most of 3G applications were considered to have much heavier traffic in the downlink direction than in the uplink direction. However, as new services such as video telephony and FTP upload are introduced and a new uploading pattern such as UCC emerged, uplink enhancements have also received great attention making High Speed Uplink Packet Access (HSUPA) and CDMA 1x EvDO Rev. A system considered in 3.5G systems [8][9].

The CDMA 1x EvDO Rev. A system was standardized in March 2004. The peak rate can achieve up to 3.1Mbps for the downlink and 1.8Mbps for the uplink. A novel flow-centric protocol design is adopted to enhance the system's capability for satisfying different flow QoS requirements. Improvement has been observed in supporting delay-sensitive applications and providing tradeoff in delay, capacity and physical-layer error-rate [7].

HSUPA was introduced to improve the capacity of WCDMA uplink in 3GPP Release 6 with the first specification version in December 2004. Although HSUPA is the commonly used terminology, Enhanced-Dedicated Channel (E-DCH) is the official term used by 3GPP to describe the new uplink transport channel [6]. The E-DCH supports fast Node B based uplink scheduling, fast physical layer hybrid ARQ (HARQ) retransmission schemes and, optionally, a shorter transmission time interval (TTI) (2ms) to reduce delays, increase the data rate (up to 5Mbps) and improve the capacity of the uplink [10].

WiMAX 802.16 has become the most promising broadband wireless access technology. In Oct 2007, WiMAX is officially accepted as the one of the 3G standards by International Telecommunication Union (ITU). The standard of WiMAX evolved from the original 802.16 to the latest 802.16e which supports full mobility. When the Orthogonal Frequency-Division Multiple Access (OFDMA) physical layer is employed, the theoretical uplink or downlink raw bitrate could achieve 70Mbps [11].

Three systems, EvDO Rev. A, HSUPA, and WiMAX, represent the most typical uplink enhanced mobile systems. Through the analysis of live traffic traces of uplink traffic in the current WCDMA network, we can suggest possible research issues for those upcoming systems. To achieve this, WiMAX module in OPNET simulator was used to show the impact of the new uplink traffic characteristics.

III. Traffic Measurement and Self-Similarity Analysis

Recent empirical studies of traffic measurements from a variety of different packet networks have demonstrated that the self-similarity or burstiness over a wide range of time scales is a prevalent phenomenon [12][13][14]. Most of network traffic measurements have been performed on wired networks with some performed on wireless downlink data networks. Little attention has been paid to uplink traffic until recently. We were only able to find one report [2] in the literature that deals with real uplink traffic obtained from WAP services in a CDMA1x network. As a growing number of end users are embracing the full potential of mobile technologies to create and share multimedia contents, most of wireless mobile operators are providing multimedia services commonly called multimedia messaging (MMS) services to support uploading of multimedia files such as pictures, video files, and music files. We have analyzed live traffic traces of MMS services collected from WCDMA networks of SK Telecom.

Currently SK Telecom has over 1,000 base stations for each network, two Serving GPRS Supporting Nodes (SGSNs) covering WCDMA networks and two Packet Data Serving Nodes (PSDNs) serving CDMA networks. Two traffic collectors are located at one SGSN and one PDSN, respectively. MMS traffic data were collected from WCDMA network for every 24 hours from August 10 to 15, 2007. In the following, we proceed to discuss the self-similarity and our analysis of this traffic trace.

There are many different definitions of self-similarity. One common definition is for continuous time processes, which states that $Y(t)$ is self-similar with self-similarity parameter H $(0 < H < 1)$ if for all $a > 0$ and $t \geq 0$,

$$Y(t) =_d a^{-H} Y(at) \quad (1)$$

where the equality is in the sense of finite dimensional distributions. A canonical example of such a self-similar process is fractional Brownian motion. But we need a definition that is more applicable to analyzing network traffic traces to estimate the self-similarity parameter or the Hurst parameter H.

Let $X(t)$, $t \in \mathbb{Z}$ be a covariance stationary stochastic process with autocorrelation function $r(k)$. Define the aggregated process $X^{(m)}$ of X at aggregation level m by averaging the original process X over a non-overlapping blocks of size m. Let $r^{(m)}(k)$ denote the autocorrelation function of $X^{(m)}$. The following is a definition of self-similarity for discrete time stochastic process.

Definition 1: $X(t)$ is exactly second-order self-similar with Hurst parameter $(1/2 < H < 1)$ if

$$r(k) = r^{(m)}(k) = \frac{1}{2}((k+1)^{2H} - 2k^{2H} + (k-1)^{2H})$$

for $k \geq 1$. $X(t)$ is asymptocally second-order self-similar if

$$\lim_{m \to \infty} r^{(m)}(k) = \frac{1}{2}((k+1)^{2H} - 2k^{2H} + (k-1)^{2H}).$$

The self-similar processes can be characterized by (i) the variance of the sample mean that decreases more slowly than the rate $m^{-\beta}$, i.e., var$(X^{(m)}) \sim c_1 m^{2H-2}$ for $1/2 < H < 1$, (ii) the autocorrelation function $r(k)$ that asymptotically behaves like $c_2 k^{2H-2}$ for $1/2 < H < 1$, and thus $\sum_k r(k) = \infty$ (long range dependence), and (iii) the spectral density $f(\cdot)$ that obeys a power law near the origin, i.e., $f(\lambda) \sim c_3 \lambda^{1-2H}$ for $1/2 < H < 1$.

These properties of self-similar processes lead to various methods for estimating the Hurst parameter H (see [15]). We estimated the Hurst parameter for MMS traffic using an analysis tool SELFIS [16]. Table I shows the results for estimated Hurst values obtained using six different methods: Aggregate variance,R/S, Periodogram, Absolute moment, Whittle estimator, and Abry-Veitch. The results demonstrate that the Hurst values are between 0.5 and 1 indicating the self-similarity.

A. Heavy-tailedness

The heavy-tailedness is roughly said to cause a network traffic to possess the self-similar property. A random variable

TABLE I
HURST PARAMETER

Estimation method	MMS (Hurst parameter)
Aggregate variance	0.920
R/S	0.739
Periodogram	0.785
Absolute moment	0.766
Whittle estimator	0.845
Abry-veitch estimator	0.745

Z is said to obey a heavy-tailed distribution if

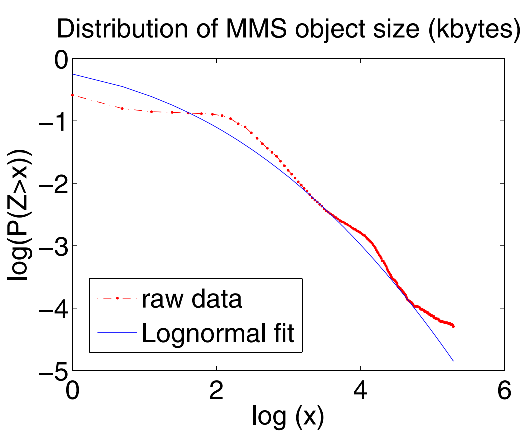
$$P(Z > x) \sim cx^{-\alpha}, \quad x \to \infty \tag{2}$$

where $0 < \alpha < 2$ is the tail index and c is a positive constant. The variable Z could represent a session size, or inter-session time between two successive sessions. In order to check for the heavy-tailedness of the distribution of a given variable, we make use of complementary distribution plots. Indeed, if we take logarithms of both sides of Eq. (2), we obtain

$$\log(P(Z > x)) \sim -\alpha \log(x) + \log(c), \quad x \to \infty.$$

This relation says that if a variable obeyed the heavy-tailed distribution, the log-log plot of the complementary distribution of the variable would be a straight line for large x-values, with a slope $-\alpha$.

In Fig. 1, we show the log-log plot of the complementary distribution of the MMS raw data. This figure suggests a strong empirical evidence of the heavy-tailed property of the file size distribution for MMS services.



(a) Log-log plot of MMS object size.

Fig. 1. Log-log plot of object size.

IV. IMPACT ON NETWORK PERFORMANCE

In this section, we analyze the impact of uplink traffic characteristics on the network performance using the WiMAX module available by OPNET 12.0 developed based on IEEE 802.16e. In our simulations, we use the MMS uplink traces collected from the SK Telecom network to model best-effort multi-media uplink traffic. Video-telephony traffic modeled in OPNET 12.0 is used to model delay-sensitive uplink traffic. The topology of a simple network is used where subscriber stations are connected to a single base station. We only model the wireless connection between BS and SSs, i.e., the radio layer in the access network, and a round-robin scheduling QoS algorithm for WiMAX available in the OPNET 12.0 is applied. The round-robin WiMAX QoS scheduling algorithm implemented in OPNET is briefly described next.

The latest IEEE 802.16e supports five scheduling service types:

- Unsolicited Grant Service (UGS)
- Extended Real-time Polling Service ($ertPS$)
- Real-time Polling Service ($rtPS$)
- Non Real-time Polling Service ($nrtPS$)
- Best Effort (BE)

UGS and $ertPS$ have the highest scheduling priority, and packets from SSs with these services are periodically inform their queue status information to the scheduler which resides in the base station. $rtPS$ and $nrtPS$ have the middle priority, and packets from SSs with these services are scheduled for transmission after clearing packets in the buffer with UGS and $ertPS$ services. BE has the lowest priority and all connections from SSs share a same scheduling queue by first-come-first-served. In our simulations, video-telephony is categorized as $rtPS$, and MMS is categorized as BE, and only traffic video-telephony services is considered delay-sensitive.

A. *Delay Analysis with Low Load Delay-Sensitive Traffic*

We first consider a scenario in which both MMS and video-telephony traffic loads are low. The parameters used for traffic generation are summarized in Table II.

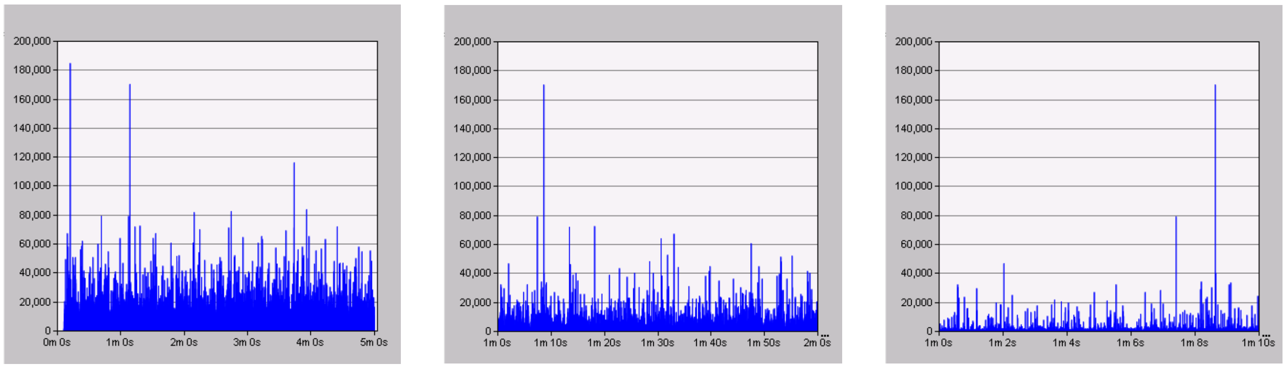
TABLE II
SUMMARY OF PARAMETERS

Parameters	Values
Simulation Time	300 seconds
Number of video-telephony SS	64
Number of MMS SS	64
video-telephony packet size (bytes)	Lognormal(6.0, 1.0)
video-telephony packet interarrival time (seconds)	Constant(0.1)
MMS packet size (kbytes)	Lognormal(1.2, 1.5)
MMS packet interarrival time (seconds)	Exponential(2.88)

Figure 2 (a) shows the total amount of aggregate traffic (in bytes/sec), over the entire simulation time, received at the server. This figure is then zoomed into 60 seconds (b) and 10 seconds (c) of simulation periods. The traffic burstiness is observed across all time scales. This scale-invariant burstiness shows the self-similarity property. Similar results are also discovered for MMS traffic and video-telephony traffic at each SS.

Since video-telephony is delay-sensitive, a major QoS requirement should be the delay. We assume that the delay requirement for this service is to have at least 99% of packets experience delays less than 0.03 seconds, and services meeting this requirement are considered satisfiable.

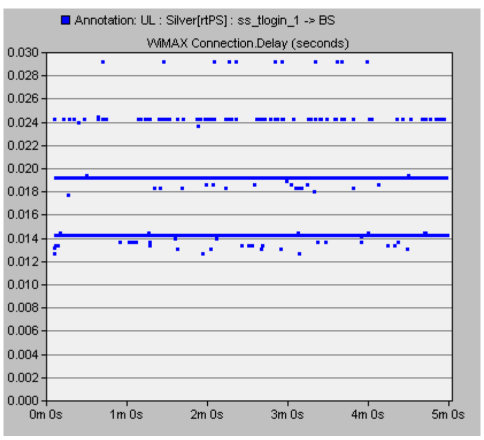
Packet delays of video-telephony service at a single SS is shown in Fig. 3(a). The QoS requirement of the video-telephony is fulfilled as shown, and it can be also clearly



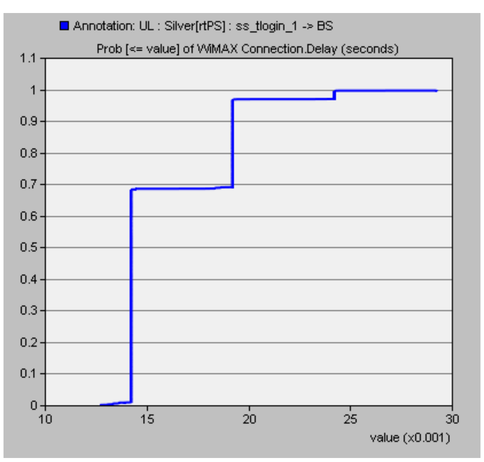
(a) 300s period (b) 60s period (c) 10s period

Fig. 2. Aggregate traffic with 64 MMS and 64 video-telephny services

verified by the cumulative density function CDF of the delay shown in Fig. 3(b).



(a) Packet delay



(b) CDF of packet delay

Fig. 3. Packet delay of video-telephony with 64 MMS and 64 video-telephony SS

B. Delay Analysis with High Load Delay-Sensitive Traffic

In this scenario, we increase the traffic loads from both MMS and video-telephony services by injecting 108 SSs with MMS services and 108 SSs with video-telephony services.

The aggregate traffic received at the BS is shown in Figure 4. We note here that the burstiness is also observed and the total traffic received at the BS (in the range of 40 to 60 bytes/sec) is much lower than the network capacity (theoretical max capacity of WiMAX is 70 Mbps).

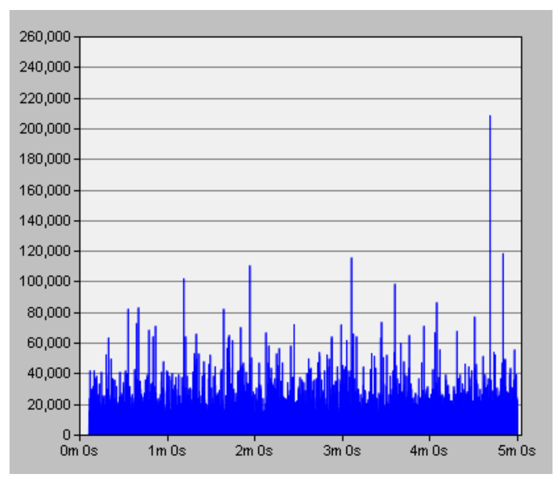
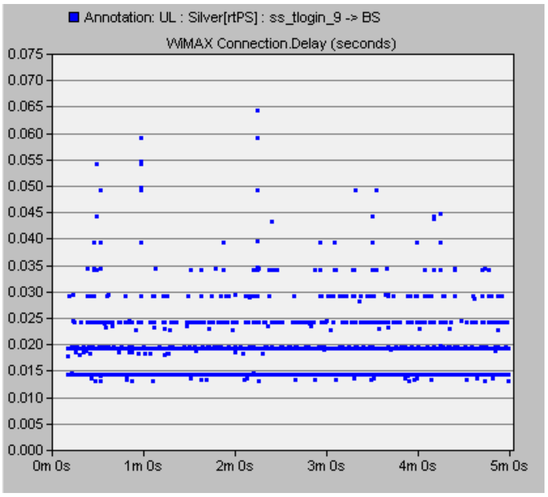
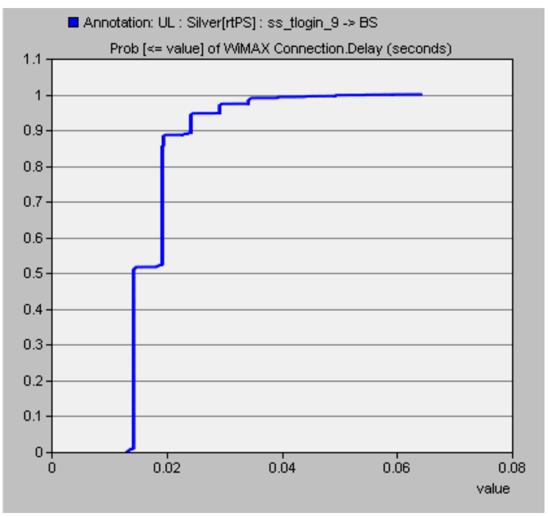


Fig. 4. Aggregate traffic with 108 MMS and 108 video-conferencing.

As the total number of SSs is increased to have 108 MMS SSs and 108 video-telephony SSs, the delay of video-telephony packets is largely affected even though they are served under the rtPS category with the higher scheduling priority. From the Fig.5(a), it is clear that the delays are much larger in this case than those in the previous scenario with 64 MMS SSs and 64 video-telephony SSs. A closer look at the CDF plot of the packet delay shown in Figigure 5(b) reveals that at least 5% packets have delays larger than 0.03 seconds. Here we should note that the aggregated traffic load in any moment is much less than 250 kbytes, which is far less than the network capacity. This observation depicts that the network resource has not been fully utilized. Better scheduling algorithms and access control schemes should be developed to improve the network's performance in supporting QoS requirements.



(a) Packet delay



(b) CDF of packet delay

Fig. 5. Packet delay of video-telephony with 108 MMS and 108 video-telephony SS

V. Conclusions

In this paper, we have analyzed uplink traffic collected from MMS services in WCDMA networks of SK Telecom, Korea. Our preliminary results using six different self-similarity analysis algorithms suggest that this uplink traffic is self-similar. We used trace-driven OPNET simulations using WiMAX module and demonstrated the burstiness in both aggregated traffic occurred at the BS and individual traffic generated at each SS. And the self-similarity is observed in a sense that and burstiness is preserved across different time scales. Our simulation results also show that the data burst from delay-sensitive uplink services can have significant impacts on the network performance even if the overall network load is lower than the network capacity. This observation suggests that call admission control should be carefully monitored for delay-sensitive uplink services as most of admission control algorithms only focus on the available network capacity. And further research on developing uplink scheduling algorithms integrated with call admission control is also required to deal with the new uplink traffic patterns.

Acknowledgment

This work was in part supported by the Access Network & Mobile Terminal R&D Center, SK Telecom, Korea. We also acknowledge OPNET Inc. for the access to their simulation tools.

References

[1] Organisation for Economic Co-operation and Development (OECD), "Participative Web and User-created Content: Wbe 2.0, Wikis and Social Networking", [Online] Available: http://www.oecd.org, 2007
[2] Liang Peng, Tang Cailin, Ma Jie, Chang Yongyu, and Yang Dacheng, "Experimental study on traffic model of wireless Internet services in CDMA network," *IEEE Proceedings of the 61st Vehicular Technology Conference*, Vol. 4, pp.96-100, April 2005
[3] OPNET, Opnet Network Modeler and Simulator, Version 12.0, [Online] Available: http://www.opnet.com
[4] 3GPP, The 3rd Generation Partnership Project, [Online] Available: http://www.3gpp.org
[5] 3GPP2, The 3rd Generation Partnership Project 2, [Online] Available: http://www.3gpp2.org
[6] H. Holma and A. Toskala,Eds., *HSDPA/HSUPA for UMTS: High Speed Radio Access for Mobile Communications*, John Wiley & Sons, Apr. 2006.
[7] Mingxi Fan, D.Ghosh, N. Bhushan, R. Attar, C. Lott, and J. Au, " On the Reverse Link Performance of cdma2000 1xEVDO Revision A System," *IEEE International Conference on Communications, ICC 2005*, Vol. 4, pp. 2244-2250, May 2005.
[8] G. Sharma and G.S. Kumar, " Moving towards HSUPA (high speed uplink packet access): a complete 3.5G wireless system," *IEEE International Conference on Personal Wireless Communications, ICPWC 2005*, pp. 174-177, Jan. 2005.
[9] 3GPP2, "cdma2000 High Rate Data Air Interface Specification - Revision A," 3GPP2 C.S20024-A, v.1.0, Mar. 2004.
[10] S.Parkvall, E.Englund, K.W. Helmersson, and M.Samuellson, " WCDMA Uplink Enhancements for High-Speed Data Access," *IEEE Proceedings of the 60th Vehicular Technology Conference, VTC2004-Fall*, Vol. 7, pp. 5204-5208, Sept. 2004.
[11] Loutfi Nuaymi, *WiMAX: Technology for Broadband Wireless Access*, John Wiley & Sons, 2007.
[12] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, " On the self-similar nature of ethernet traffic (extend version)," *IEEE/ACM Trans. Networking*, vol.2, no.1, pp.1.15, Feb. 1994.
[13] M. E. Crovella and A. Bestavros, "Self-similarity in world wide web traffic: evidence and possible causes," *IEEE/ACM Trans. Networking*, vol.5, no.6, pp.835.846, Dec. 1997.
[14] M. Jiang, M. Nikolic,S. Hardy, and L.Trajkovic, " Impact of self-similarity on wireless data network performance," *IEEE International Conference on Communications, ICC 2001*, Vol.2, pp. 477-481, June 2001.
[15] A. Popescu, " Traffic Self-Similarity," *IEEE International Conference on Telecommunications, ICT2001*, Bucharest, June 2001.
[16] T. Karagiannis, M. Faloutsos, " SELFIS: A Tool For Self-Similarity and Long-Range Dependence Analysis," 1st Workshop on Fractals and Self-Similarity in Data Mining: Issues and Approaches (in KDD), Edmonton, Canada, July 23, 2002.