End-to-end Wireless Performance Simulator : Modelling Methodology and Performance

Sung-Min Oh, Hyun-Jin Lee and Jae-Hyun Kim

School of Electrical and Computer Engineering
Ajou University, Suwon, Korea
{jkim, l33hyun, smallb01@ajou.ac.kr}

Abstract. To predict application level performance in wireless networks, we build wireless performance simulator including application traffic characteristic, network architecture, network element details and protocol features. we also develop the simulation modelling methodology using Lindley's recursion method to reduce the number of simulation events. Using the simulator, we assess the user perceived application performance of the voice and web browsing service in the cdma2000 network for the wireless technology migration from 2.5G to 3G+. The main conclusion is that end-to-end application-level performance is affected by various elements and layers of the network and thus it must be considered in all phases of the development process.

1 Introduction

cdma2000 3G-1X RTT (Radio Transmission Technology) was on the market from 2001. Many wireless service providers have been considering the 3G wireless technology migration path from circuit to packet technologies. Since the ATM transport can be used to support QoS in current technology but technology migration trends are looking for All IP transport in near future. User performance studies for cdma2000 were published in many papers [1–4], In [1], data service performance was evaluated for 3G-1X RTT system but an alternative architecture or voice service was not addressed. In [2], the TCP performance was presented in wireless interface but an end-to-end performance was not included. Most of the papers addressed the wireless channel throughput or sector throughput and some others considered QoS strategies in cdma2000 [3, 4]. However, very few studies considered the whole network architecture. The user perceived application performance should be considered in an end-to-end reference architecture including a Radio Access Network (RAN), a Core Network (CN) and a data center, otherwise we can get only partial information on the application-level performance. To assess the user perceived application-level performance characteristics of different QoS service classes for alterative transport technologies and wireless technology evolution scenarios, we propose an end-to-end performance simulator for 2.5G or 3G+ networks. In this paper, we describe the end-to-end performance simulation model and methodology that we built for cdma2000 network. We model all the protocol layers from the physical through the application layer and model details of the packet handling characteristics of each network element along the path. Foreground and background traffic loads are generated to represent specific application environment. We also address application-level performance issues in terms of wireless technologies evolution from 3G-1X RTT to 3G-1X EV and transport technology evolution from ATM to IP. In general, the goal of the simulator is to design quality and performance into the products and solutions up front. Some highlights of the wireless performance simulator are:

- Model based on the OPNET simulation tool
- Models end-to-end reference connections
- Models cdma2000 applications and services
- Models user-plane traffic
- Impacts of mobility will be approximated
- Models each network element
- Models cdma2000 packet flow and detailed protocol stacks
- Models QoS buffering and scheduling algorithms for network elements

2 Network Simulation Models

2.1 Reference architecture and connection models

We study the performance modelling for the 2.5G and 3G+ networks. The 3G-1X system supports data rate from 9.6 kbps to 2.4 Mbps[5]. Fig. 1 shows a reference network architecture model for the cdma2000. The reference network architecture can be considered into four different networks; RAN , CN, internet and data center. RAN may include Mobile Terminal (MT), Base Station Transmission System (BTS), Base Station Controller (BSC), Mobile Switching Center (MSC) and ATM or IP concentrators. CN includes ATM or IP routers and Packet Data Serving Nodes (PDSN). The data center network can be composed of three zones to protect servers from hacking or virus; a public, a DMZ (Demilitarized Zone), and a secure zone. Each zone can be protected by firewalls as shown in the Fig. 1.

2.2 Protocol architectures and models

In this paper, we consider two transport technologies such as ATM and IP in the RAN and CN. For ATM transport scenarios, a BTS chops a reverse link traffic packet in to ATM cells and transmits them to MSC or Radio Network Controller (RNC) (for ALL IP scenario). Voice traffic uses AAL2 and data traffic uses AAL5 layers respectively in RAN. For IP transport scenarios, BTS transmits a IP packet on the top of T1 and IP router converts it to Ethernet packet and sends it to MSC. The detailed protocol stack for IP protocol architectures are shown in Fig. 2. To assess the application-level performance we implement all the protocol stacks in Fig. 2 except wireless channel model. To simulate the wireless channel error, we used the following link level simulation results. The channel

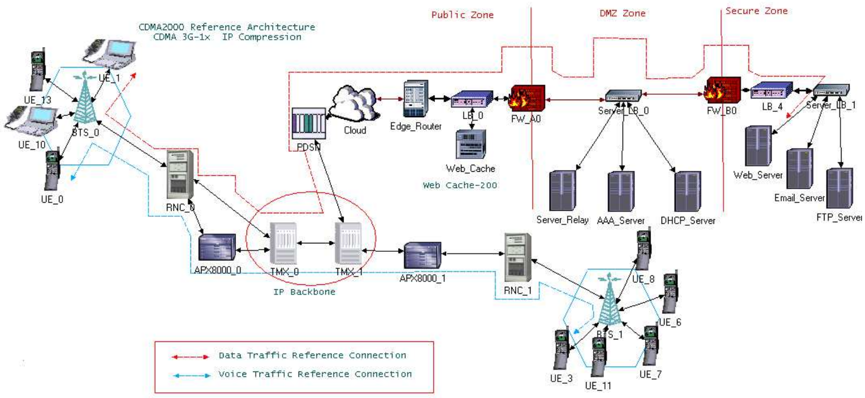


Fig. 1. Reference network model for cdma2000

model used in this paper is based on the models specified in 3G 1X-RTT. For link level simulation, we use a traced file which contains frame error data when the target frame error rate is fixed at 1%, 4% or 10%. These error data are time co-related for each frame upon channel model.

3 Service Traffic Models

We build two types of traffic model in the simulator: foreground and background traffic. The foreground traffic represents the specific traffic and services that traverse a given reference connection. It is the goal of the simulator to model the

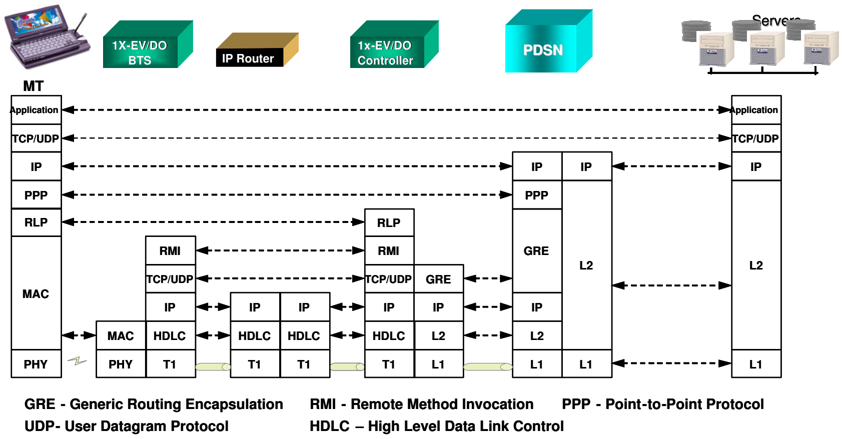
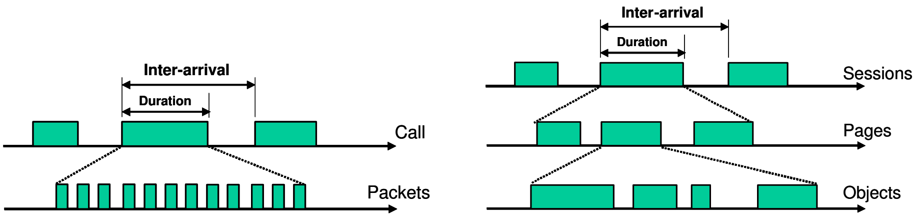


Fig. 2. Protocol stack model for All IP CDMA 2000



(a) Voice traffic model (b) Web browsing traffic model

Fig. 3. the hierarchical structure for voice and web services

performance of this traffic in detail. An important contributor to the performance of these services will be the characteristic and load contributed by the background traffic at each node.

3.1 Foreground traffic load models

We use a voice traffic model and a web browsing traffic model for the applications in the paper. The voice traffic is generated by two hierarchical structures; call and packet level. Call level model composes of a sequence of ON and OFF periods as shown in Fig. 3(a). Each duration of ON and OFF period are exponentially distributed with mean 3 sec. (activity factor is 0.5). During the ON period MT generates Enhanced Variable Rate Codec (EVRC) 8 kbps voice traffic packets[6]. We use the 3GPP2 standard traffic model for a web browsing service[7]. An example of the design model for a web browsing service is illustrated in Fig. 3(b). We characterize the arrival of page requests within a session, and the number of objects and their sizes for each page. Detailed statistics can be found in Table 1. Other applications of interest can be modelled similarly. It is the performances of these foreground applications that the simulation model will measure in detail.

3.2 Background traffic load models

The background traffic load must be representative of the applications that are being run over the network so that its impact on the foreground traffic is accurately accounted for. However, the impact on the simulation run time precludes detailed application-level models. After reviewing an enormous number of different methods, such as statistical models for data traffic (long range dependant type)[8] and traffic analysis and synthesis[9], we decide to use a traffic trace to get the effect of the background traffic load. The first step is to collect a detailed packet trace for 1000 simultaneous application sessions using the simulator. This trace file is then scaled to match the desired mean rate for a given application. The trace file approach improved simulation run-time performance, but it was still too slow to run large scale network simulations. Thus we used the trace file to simulate a virtual packet load by calculating the delay effect in the buffer instead of generating background traffic packet by packet. To calculate the packet delay effect we used Lindley's recursion method and extended it to account for the impact of multiple queues and queue scheduling disciplines. Lindley's recursion equation is given by

$$W_q^{(n)} = \begin{cases} W_q^{(n-1)} + S^{(n-1)} - T^{(n-1)} & , \ if \quad \left(W_q^{(n)} + S^{(n)} - T^{(n)} > 0\right) \\ 0 & , \ otherwise \end{cases} \tag{1}$$

where, $W_q^{(n)}$ and $W_q^{(n-1)}$ mean waiting times of the n^{th} packet and $(n-1)^{th}$ packet respectively. $S^{(n-1)}$ denotes the service time of the $(n-1)^{th}$ packet and $T^{(n-1)}$ means the inter-arrival time between the $(n-1)^{th}$ and n^{th} packets. The packet delay calculation algorithm for multiple priority queue is as following.

- Definition
  - F_0, F_1, $\cdots$, F_i, $\cdots$, F_p: Background trace file with priority 0, 1, $\cdots$,i, $\cdots$, p. 0 is highest priority and the F_i is the traced file for the current reference packet with priority i.
  - *t_last*: the time that the $(n-1)^{th}$ reference packet is arrived
  - *t_arv*: the time that a n^{th} reference packet arrived
  - *t_ia*: inter-arrival time between the $(n-1)^{th}$ packet and n^{th} reference packet
  - *t_wait*: waiting time for the n^{th} reference packet which calculated by Lindley equation
  - *t_serv*: the service time for the $(n-1)^{th}$ packet for the n^{th} reference packet
  - *t_dep*: the departure time for the n^{th} reference packet. *t_dep* = *t_arv* + *t_wait*
- Algorithm

step 1. Calculate waiting time for the reference packet (priority i) for the F_i.

```
while  (t_last + t_ia <= t_arv) {
    t_wait = t_wait + t_serv - t_ia ;
    if (t_wait < 0) {
        t_wait = 0;}
    t_last = t_last + t_ia;}
```

step 2. If there are any packets between *t_arv* and *t_dep* in the any of higher priority background trace file, then repeat step 1 until no other higher priority traced packet are between *t_arv* and *t_dep*.

step 3. If there is any higher packet(s) arrived between *t_arv* and *t_dep*, defer the *t_dep* by the service time of the higher priority packet(s) and recalculate *t_dep*

step 4. Repeat step 2 and 3 until there is no any other reference of background packet between *t_arv* and *t_dep*

4 Simulation Model Description

4.1 Network architecture model description

In this study, we consider voice and data service scenarios. To evaluate application-level performance with different network configuration, we consider the following five scenarios:

- Scenario 1 (2.5G Tandem switch): A voice packet is initiated from MT and transferred to BTS, the BTS then sends the voice packet on the ATM/AAL2 to MSC. In the MSC, Interworking Function (IWF) converts an EVRC packet to PCM 64 kbps packet format and sends it to tandem switch.
- Scenario 2 (3G ATM, G.711): From MT to MSC is the same as Scenario 1 but IWF in MSC converts an EVRC packet to a PCM packet format and sends it to a media gateway. The media gateway transfers PCM packet to ATM CN using ATM/AAL1
- Scenario 3 (3G ATM, G.726): From MT to media gateway is the same as Scenario 2. The media gateway converts a PCM 64 kbps packet to a G.726 ADPCM 32 kbps packet and sends it to ATM core network using AAL2 multiplexing.
- Scenario 4 (3G+ IP, VoIP): MSC converts an EVRC packet to a G.726 32 kbps packet and sends it to IP based media gateway using 100BT Ethernet. Then the IP media gateway sends it to IP CN.
- Scenario 5 (3G+ All IP, VoIP, vocoder bypass): This is All IP scenario. IP based BTS sends EVRC packet to IP RNC using 100BT Ethernet. The RNC then transfers the EVRC voice payload over an IP packet to IP CN.

4.2 Network elements and physical layer model description

Two firewalls, two load balancers and 3 routers are modeled in a data center. For CN elements, media gateway, ATM switch and IP router models are implemented. We also implemented MSC(for 2.5G and 3G), RNC(for 3G+), BTS and MT models for RAN elements. The packet processing time for each network element follows the 3GPP standard specification [10]. We fully implemented each protocol in the network elements shown in Fig. 2. The air channel and physical layer is modeled based on the average channel quality and mobility. The user mobility model assumed that mobile users are uniformly distributed in a cell. Mobile users are assumed to move at a pedestrian speed of 3 km/h with worst case fading of single path Rayleigh. Based on the location of mobile terminal, we use the link level simulation result to estimate the power requirement for the user requested data rate. we have implemented a 3G 1X-RTT packet scheduler and a proportional fair scheduling algorithm for 3G 1X-EV scenario based on [11]. Some of simulation parameters are summarized in Table 1.

5 Simulation Results and Discussions

In this section, we assess the performance of CDMA 2000 systems in terms of end-to-end delay for voice and data traffic service for 2.5G, 3G 1X-RTT and 3G 1X-EV.

5.1 Voice service performance simulation results

To compare an end-to-end voice packet delay performance for technology evolution from 2.5G to 3G+, we perform the simulation for the five different scenarios.

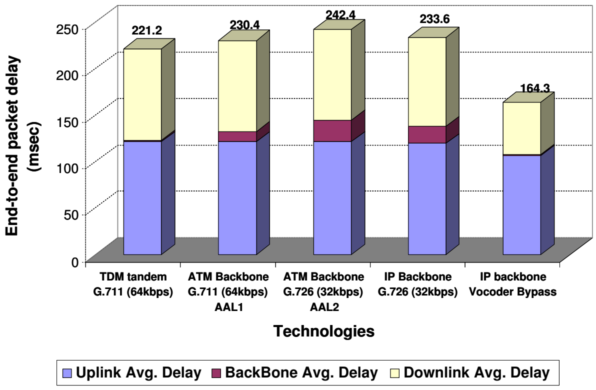
Table 1. Simulation Parameter

Category	Parameter	Reference
Voice traffic	EVRC 8 kbps	[6]
Web browsing traffic	Main object size: lognormal(10.8,250)Kbyte Embedded object size: lognormal (7.8,126) Kbyte Number of objects per page: Pareto shape :1.1, location: 55	[7]
TCP parameters	Windows 2000 based parameters	[7]
Radio Link Data Rate (kbps)	9.6, 153.6, 2000, 2400	[5]
RLP scheme	(2,3) RLP scheme	[2]
Processing time (msec)	MT - forward : 36.55, reverse : 63.05 BTS - forward : 15 ,reverse : 9 MSC/RNC - forward : 7 ,reverse : 7 ATM/IP router: 0.1 ,Internet : 1. IP router processing time:100 μsec	[10]

The results are presented in Fig. 4. The background traffic load for each network element is 40% in the simulation. In scenario 2 and 3 (ATM CN), voice packet delay is a little bit larger than that of the scenario 1 (tandem switch). Because in both scenarios ATM processing delay is included in the CN, and the AAL2 multiplexing delay is (processing delay and Timer_CU : 2msec) also included in CN for the scenario 3. In the scenario 4 (IP CN, G.726 ADPCM), the CN packet delay is increased compare to it in the scenario 2 (ATM AAL1) since the IP packet overhead for EVRC voice traffic is larger than packet format overhead of ATM. The scenario 5 is for the vocoder bypass which means that an EVRC voice packet is transferred over the IP packet without any transcoding to other coding scheme. It reduces RAN and CN coding processing delay and results into 30% delay reduction compared to the scenario 3. If we map these one-way end-to-end delay to the R value in E model in [12], the voice quality for the scenario 1 and 2 provide high quality voice. The scenario 3 and 4 provide medium quality voice, while the vocoder bypass scenario meets the high quality voice.

5.2 Data service performance simulation results

ATM vs. IP in Radio Access Network ATM transport technology in current 3G network will eventually migrate to IP technologies. Fig. 5(a) presents the web page response time for three different RAN transport technologies: ATM, HDLC over T1, and 100BT Ethernet. At 10% FER, we observe 15.6% and 21.7% page response time reduction when RAN transport technology migrates from ATM to HDLC and ATM to 100 BT Ethernet, respectively. The 21.7% performance improvement is due to the higher transmission speed and lower packet overhead in IP layer. In this case, IP transport technology is better solution for higher FER environment since the IP packet overhead is smaller than that of ATM

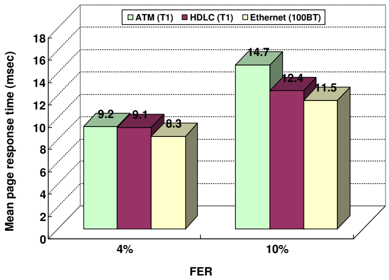


Voice Quality Scores (R value)

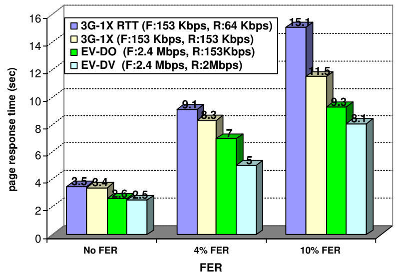
E2E One Way Delay (ms)	64 kbps (G.711)	32 kbps(G.726) 16kbps(G.728)
0	94	87
50	93	86
100	92	85
150	90	83
200	87	80
250	80	73
300	74	67
350	68	61
400	63	56
450	59	52

90 - 100	Best quality; users very satisfied
80 -90	High quality; users satisfied
70 -80	Medium quality; some users dissatisfied
60 -70	Low quality; many users dissatisfied
50 -60	Poor quality; nearly all users dissatisfied
Less than 50	Unacceptable quali ty

Fig. 4. end-to-end voice packet delays for technology evolution



(a) different RAN transport technologies



(b) 3G technology

Fig. 5. Web page response time for different RAN transport and 3G technology

for web browsing data traffic, while it shows opposite effect to small size voice packet as shown in Fig. 4.

3G-1X RTT vs. 3G-1X EV 3G-1X EV service started from 2001 in Korea. 3G-1X EV enhances the data rate to 2.4 Mbps. To compare the data service performance for 3G wireless technology, we measure the web browsing response time for different data rates in 3G-1X RTT and EV networks. We assume that 100 BT Ethernet RAN and IP CN transport technology in this scenario. There is no significant performance difference when the channel is error free. However, at 10% FER, EV provides 46% reduced response time compared to 1X RTT. The higher data rate in RAN of EV results in faster frame retransmission compared to 1X RTT. As FER increase, the performance difference between the two technologies becomes more significant.

5.3 Simulation Runtime Performance

The run-time performance of the simulation can be defined in terms of number of events and processing time per event. The simulation run-time performance is always an important issue, but is especially so for network simulations where the number of events can be extremely large. As mentioned previously, we have separated the traffic into foreground and background traffic and developed specialized techniques for handling each to improve the simulation efficiency. To quantify the wireless performance simulator, we modelled FTP applications with varying numbers of users: The FTP application was a 1 Mbyte file download over the 64 kbps data rate and Table 2 shows some of the simulation run times. The simulation takes 120 sec for a single user and file and the simulation time increased linearly when the number of concurrent FTP sessions were added. It clearly shows that the simulation performance is not feasible when the number of concurrent application sessions is large. However, the last two rows of Table 2 show that the simulation performance is improved when additional FTP sessions are modelled as background traffic. For this scenario, 124 Mbps and 147 Mbps traffic on the average, which is about 80% and 95% of STM-1, was generated in all of the nodes along the reference connection (excluding application server) and one foreground FTP session was created.

6 Conclusions

In this paper, we described an end-to-end performance simulation model and methodology that we had built for cdma2000 network. The simulator modelled all protocol layers from the physical through the application layer and modelled details of the packet handling characteristics of each network element along the path. We addressed application level performance issues in terms of wireless technologies evolution from 2.5G to 3G+. We found end-to-end QoS mechanism should be provided in every network elements where the packet passes by. The main contributions of this paper are threefold:

1. Develop the new simulation methodology using trace file and Lindley's recursion method to improve the simulation runtime performance

Table 2. Simulation Run Time with and without Background traffic model

Number of Foreground Users	Number of background Users	Download File size	Simulation Time (sec)
1	0	1 Mbytes	120 sec
2	0	1 Mbytes	237 sec
3	0	1 Mbytes	355 sec
4	0	1 Mbytes	478 sec
5	0	1 Mbytes	596 sec
1	1400	1 Mbytes	190 sec
1	1670	1 Mbytes	205 sec

2. Build end-to-end network simulation model for cdma2000
3. Access user perceived application performance for voice and data services

Wireless performance simulator presented in this paper has been used to predict and quantify the performance of cdma2000 applications, services, and network architectures.

References

1. Z. Dziong, F. Khan, K. Medepalli, S. Nanda, "Wireless Internet Access Using IS-2000 Third Generation System: A Performance and Capacity Study," *Wireless Networks*, vol. 8, pp. 325–336, 2002
2. E. Chaponniere, S. Kandukuri and W. Hamdy, "Effect of physical layer bandwidth variation on TCP performance in CDMA2000," *IEEE Vehicular Technology Conference spring 2003*, Jeju, Korea.
3. F. Li, M. Nguyen and W. Seah , "QoS Support in IP/MPLE-based Radio Access Networks," *IEEE Vehicular Technology Conference spring 2003*, Jeju, Korea
4. O. Sallent, J. Romero, R. Agusti and F. Casadevall, "Provisioning Multimedia Wireless Networks for Better QoS: RRM Strategies for 3G W-CDMA," *IEEE Commun. Mag.* pp. 100-106, Feb, 2003.
5. 3GPP2 S.R 0023 V.20, "High-Speed Data Enhancements for cdma2000 1x Data Only" 2000-11
6. 3GPP2 C.S0014-0 V.1.0, "Enhanced Variable Rate Codec," 1999-12
7. 3GPP2 TSG-C.R1002, "1xEV-DV Evaluation Methodology (V13)," 2003
8. A. Reyes-Lecuona, et al. "A page-oriented WWW traffic model for wireless system simulations,"Proceedings of ITC-16, pp 1271-1280, 1999
9. M. Lucas, B. Dempsey, D. Wrege and A. Weaver," An Efficient Self-Similar Traffic Model for Wide-Area Network Simulation," IEEE GLOBECOM '97, Phoenix, AZ, November 1997
10. 3GPP TR 25.853 v4.0.0, "Delay Budget within the Access Stratum," 2001-03
11. A.Jalali, R.Padovani and R.Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," *IEEE Vehicular Technology Conference*, Tokyo, May 2000
12. M. Perkins, C. Dvorak, B. Lerich, J. Zebarth, "Speech Transmission Performance Planning in Hybrid IP/SCN Networks," *IEEE Commun. Mag.* pp.126-131, July 1999.