Statistical Analysis of End-to-End Delay for VoIP Service in Mobile WiMAX Networks

Mohd. Noor Islam*, ***Regular Member,*** Yeong Min Jang* ***Lifelong Member***

ABSTRACT

Measurement of Quality of Service (QoS) parameters and its statistical analysis becomes a key issue for Mobile WiMAX service providers to manage the converged network efficiently and to support end-to-end QoS. In this paper, we investigate the population distribution of end-to-end one-way delay which is the most important QoS parameter in Mobile WiMAX networks. The samples are analyzed with Chi-Square Goodness-of-Fit test, Kolmogorov-Smirnov (K-S), and Anderson-Darling (A-D) test to verify the distribution of parent population. The relation with confidence level and the minimum number of sample size is also performed for logistic distribution. The statistical analysis is a promising approach for measuring the performance Mobile WiMAX networks.

Key Words : Mobile WiMAX, QoS, Logistic distribution, confidence interval (CI), Minimum sample size

Ⅰ. Introduction

The Mobile WiMAX is using the 2.3GHz frequency, and maximum transmission speed is 8 Mbps with maximum 60 km/h mobility. The Mobile WiMAX services can be classified into four categories as unsolicited grant service (UGS), real time polling service (rtPS), non-real-time polling service (nrtPS) and best effort service (BE). QoS comes into attention when planning, deploying network as well as it is important to monitor in order to support different tiers of service for different customers. Guarantees on data rate, latency, jitter, and error rate are critical as far as Mobile WiMAX is concerned; otherwise it could not compete with other similar technologies such as 3G Cellular, Wi-Fi, DSL, and Cable. And for applications such as enterprise video conference, high quality video streaming, depends on the underlying network characteristics.

QoS is important because it enables the service provider to manage the network in the most efficient manner; this is a key issue in an area like wireless which is always ruled by scarcity of spectrum and capacity. Internet Service Providers (ISPs) now offer service level agreements (SLAs) routinely to their customers. This has driven the service-providers to seek consistent testing and measurement methods to accurately measures network performance. Developing a proper monitoring and performance estimation techniques therefore becomes a key challenge for network management. However, the implementation of measurement becomes increasingly difficult and complex due to the rapid expansion of the internet Moreover, the dramatic increase in the speed of wide area backbones presents obstacles to complete statistics collection. The enormous amount of measurement data may significantly increase the cost and resource usage[1].

※ "This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute for Information Technology Advancement)" (IITA-2009-C1090-0902-0019)

* Kookmin University, Korea(yjang@kookmin.ac.kr

논문번호 : KICS2009-12-604, 접수일자 : 2009년 12월 5일, 최종논문접수일자 : 2010년 2월 10일

Sampling-based measurement methods are used to reduce the quantity of control data and resources required for network performance monitoring and finally to reduce the measurement complexity and cost[2]. Small number of samples may not represent the exact parent population and true value of parameters of the population distribution but we can get an idea about them by statistical analysis. We can make an interval for the true value of the parameters with a certain confidence level but we should consider a percentage of error. So it is also essential to know the relation with error and sample size.

One important QoS parameter of VoIP service in Mobile WiMAX network is end-to-end one-way delay. This paper will present statistical analysis about that parameter. The rest of the paper is organized as follows. In the next section we explain the data collection procedure for Mobile WiMAX networks. We analyze the sampled data to fit for distribution in Section III. Section IV provides the theory of confidence interval (CI) and sample size and the numerical analysis of sampled data is shown in Section V. Lastly we conclude our paper by Section VI.

Ⅱ. Performance Measurement Architecture for Mobile WiMAX Networks

Samples of end-to-end one-way delay for VoIP service are taken using the WiBro network in Seoul, Korea. The network architecture for performance measurement is shown in Figure 1. We used laptop as user main terminal. We used KT WiBroCARD (Model: SWT-H200K) for using WiBro network. We took samples of end-to-end one-way delay using the web site http://speed.nia.or.kr provided by National Information Society Agency (NIA), Korea. The codec was set at G.711 (Home phone) and for the measurement the setting was selected that one person is in the network. We also set special HTTP protocol for VoIP service. In Figure 1, a packet is sent from the laptop to the Measuring Server, the Measuring Server calculate the end-to-end one-way delay and it

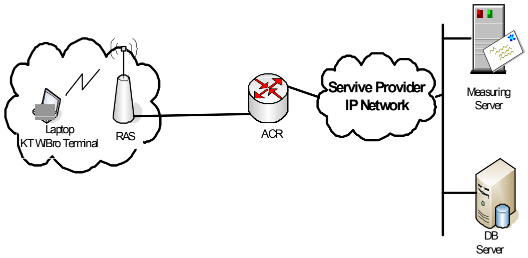

Fig. 1. Mobile WiMAX network architecture for data collection

stored the data in the Data Base(DB) Server. We collected data from the DB Server through internet.

As Mobile WiMAX is a mobile internet, we took data from different places. For stationary condition we took data from the first floor of the same building and also took data about 5m near to the RAS. A summary of sample data is shown in Table 1.

Table 1. Measurement sample data for end-to-end one-way delay

User Terminal motion condition	Different places	**Maximum value [ms]**	Average **[ms]**
Stationary Condition	1st floor: NLOS	48.24	43.40
	Near to RAS: LOS	44.77	42.66

Ⅲ. Fitting for the Measurement Data

We analyze the fitting for parent distribution of samples. We analyzed all the samples collected from different places and different conditions using three popular hypothesis tests such as Chi-Square Goodness-of-Fit test, Kolmogorov-Smirnov (K-S), and Anderson-Darling (A-D) test. We tested among exponential distribution, Pareto distribution, Weibull distribution, Weilbull (3P) distribution, Normal distribution, Student's t distribution, Chi-Square distribution. We used EasyFit 5.0 software for hypothesizes testing. We did hypothesis tests for five significant values (α) such as 0.5, 0.2, 0.05, 0.02 and 0.01 to check samples are accepted or rejected for a particular distribution. Pareto distribution is most appropriate to model end-to-end delay while normal or normal-related distribution is not for wide area network[3]. But in case of our analysis measurement with the Mobile WiMAX, end-to-end one-way delay is shown that it is better fitted with the logistic distribution. Distribution fitting analysis of samples taken from stationary condition of main user terminal is described below. We may recommend that the distribution of end-to-end one-way delay is logistic distribution. Figure 2 shows the P-P plot for logistic distribution for the same data.

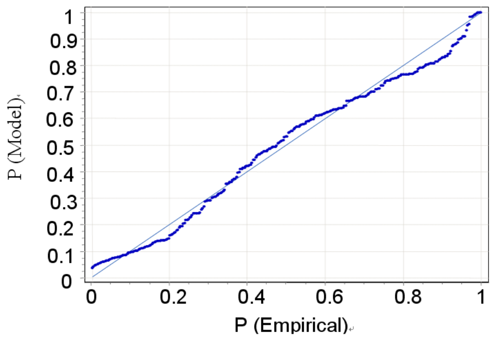

Fig. 2. P-P plot for logistic distribution

The logistic distribution has two parameters: continuous scale and location parameter. The parameters are calculated by Maximum Likelihood Estimation (MLE). The parameters for different samples to fit with the logistic distribution are shown in Table 2.

Table 2. Parameters for different samples of logistic distribution

Number of samples	Parameters	Better fitted distribution
200	$\hat{a}$ = 39.177, $\hat{b}$ =1.715	Logistic
300	$\hat{a}$ = 38.5, $\hat{b}$ =2.077	Logistic
350	$\hat{a}$ = 38.62, $\hat{b}$ = 2.08	Logistic

Ⅳ. Analyzing Confidence Interval and Determining the Minimum Sample Size

In this section we discuss on confidence interval and minimum sample size for a certain error for the parameters of the distribution of end-to-end delay. Sample size and the confidence interval of an estimate are mutually related to each other. By different hypothesis test we already conformed that end-to-end one-way delay of VoIP service in WiBro network follows logistic distribution. It has two parameters such as location parameter and scale parameter. In this section we are going to formulate mathematically the confidence interval for the both parameter as well as the equation for minimum sample size relating with the confidence interval. Recall that we used EasyFit 5.0 software to test the fitness of distribution. That software use MLE to estimate the location and scale parameters.

Let the set of random variables $x_1, x_2,..,x_n$ is called a sample of size n from the population. The sample mean is μ and sample standard deviation is σ.

$$\mu = \sum_{i=1}^{n} \frac{x_i}{n} \tag{3}$$

and

$$\sigma = \sqrt{\left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)\right)} \tag{4}$$

Let the location parameter be a and scale parameter be b and $\hat{a}$ and $\hat{b}$ are their estimators, respectively. The variance and covariance of $\hat{a}$ and $\hat{b}$ are calculated from Fisher Matrix. The variance of $\hat{a}$ and the variance of $\hat{b}$ can be written as[4]

$$Var(\hat{a}) = \frac{9\sigma^2}{n\pi^2} \tag{5}$$

$$Var(\hat{b}) = \frac{9\sigma^2}{(3+\pi^2)n} \tag{6}$$

where n is the sample size and σ is the standard deviation of the samples which can be calculated using equation (4).

4.1 Mathematical analysis of confidence interval

When the sample size is large, then the confidence level is increased and when the sample size is small, the confidence level is decreased. The confidence interval for location parameter and scale parameter for the logistic distribution are discussed below.

4.1.1 Confidence interval for location parameter

The confidence interval of the location parameter of logistic distribution can be written as [4]:

$$\hat{a} - K_{\alpha}\sqrt{Var\ (\hat{a})} \le a \le \hat{a} + K_{\alpha}\sqrt{Var\ (\hat{a})} \quad (7)$$

Putting the value of $Var\ (\hat{a})$ from (5) in (7) we get the following equation:

$$\hat{a} - K_{\alpha}\frac{3\sigma}{\sqrt{n}\,\pi} \le a \le \hat{a} + K_{\alpha}\frac{3\sigma}{\sqrt{n}\,\pi} \quad (8)$$

Equation (8) is representing the confidence interval for location parameter of logistic distribution. Estimate of location parameter, $\hat{a}$, can be calculated using MLE, σ can be calculated from the samples. K_{α} is defined by:

$$\alpha = \frac{1}{\sqrt{2\pi}}\int_{K_{\alpha}}^{\infty} e^{-\frac{t^2}{2}}dt = 1 - \phi(K_{\alpha}) \quad (9)$$

Let the confidence level, δ, then $\alpha = \frac{1-\delta}{2}$ for the two sided bounds and $\alpha = 1 - \delta$ for the one-sided bound.

From (9) it can be written that

$$\varphi(K_{\alpha}) = \frac{1}{2}\left[1 + erf\ (\frac{K_{\alpha}}{\sqrt{2}})\right] = \delta \quad (10)$$

and

$$K_{\alpha} = \sqrt{2}\left(erf^{-1}(2\delta - 1)\right) \quad (11)$$

From the equation (11), it is clear that if we decide the confidence level, then for that confidence level we can calculate K_{α}. Using the value of K_{α}, from equation (8) we can calculate the confidence interval for the location parameter.

4.1.2 Confidence interval for scale parameter

The confidence interval for the scale parameter b of logistic distribution can be written as[4]

$$\hat{b}\,e^{\frac{-K_{\alpha}\sqrt{Var\ (\hat{b})}}{\hat{b}}} \le b \le \hat{b}\,e^{\frac{K_{\alpha}\sqrt{Var\ (\hat{b})}}{\hat{b}}} \quad (12)$$

Putting the value of $Var(\hat{b})$ from equation (6) in equation (12) we get the following equation

$$\hat{b}\,e^{\frac{-K_{\alpha}3\sigma}{\hat{b}\sqrt{n(3+\pi^2)}}} \le b \le \hat{b}\,e^{\frac{K_{\alpha}3\sigma}{\hat{b}\sqrt{n(3+\pi^2)}}} \quad (13)$$

Equation (13) is representing the confidence interval of scale parameter of logistic distribution. The value $\hat{b}$ can be determined from MLE and the value of K_{α} can be measured using equation (11).

4.2 Mathematical analysis of minimum sample size

With collecting large sample size, it costs more time and memory space. So it is important to know the relation of confidence level and sample size to estimate the parameter of a distribution. Sample size can be calculated in two ways such as (i) when the population size of the distribution is unknown and (ii) the population size of the distribution is known. In general, the population size of the distribution is considered unknown, so we analyze the case for the population size, N, is unknown.

Let the standard error be E, then

$$E = K_{\alpha}\sqrt{Var\ (\hat{a})} = K_{\alpha}\frac{3\sigma}{\sqrt{n\pi^2}} \quad (14)$$

Therefore,

$$n = K_{\alpha}^{2} \frac{9\sigma^{2}}{\pi^{2}E^{2}} \quad (15)$$

From equation (15), it is seen that sample size is depends on K_{α}, i.e. confidence level, δ, and error, *E,* that can be tolerable by the system of service provider and variance of the samples taken for any service of the network. This equation represents that if the confidence level is increased the sample size needs to be increased and if the error tolerance is decreased then the sample size will be increased. So it could be find out the appropriate values for analyzing the performance statistics of the system.

Ⅴ. Numerical and Statistical Analysis of Measurement Data

We took 350 samples of end-to-end one-way delay for VoIP service from WiBro network. At first the standard deviation of those samples is calculated using equation (4). The standard deviation of the samples is 3.17. The value of K_{α} can be calculated using the equation (11). The value of K_{α} is related to the value of confidence level. Different values of K_{α} for different confidence levels are shown in Table 3 below.

Now using equation (15), we calculate the sample size needed to get a specified confidence level under a tolerable error of the system. We calculate the minimum sample size for 95% and 99% confidence level and 1% to 10% error. The minimum sample size is shown in Table 4.

From the Table 4 it is seen that if the service provider fix the tolerable error level value at low, then for a certain confidence level sample size is large and if they fix the error at high value, then the sample size becomes small. From the Table 5 they can decide which value will be appropriate for them then they can decide that value.

Table 3. K_{α} for different confidence level, δ

Confidence level, δ [%]	90%	95%	99%
K_{α}	1.28	1.64	2.33

Table 4. Minimum number of samples for end-to-end one-way delay

Error, E [%]	Minimum number of samples, *n*	
	95% confidence level	99% confidence level
1	249,033	495,916
2	62,258	123,979
3	27,670	55,101
4	15,564	30,994
5	9,961	19,836
6	6,917	13,775
7	5,082	10,120
8	3,891	7,748
9	3,074	6,122
10	2,490	4,959

We also calculate confidence interval for the location and scale parameters. Confidence interval represents an interval of values, the exact value is inside the interval. exact value of estimated parameter. The confidence interval of location parameter is calculated using equation (8) and the confidence interval for scale parameter is calculated using equation (13). Table 5 shows the calculated confidence interval for 95% confidence level and 99% confidence level. From the Table 5 it can be seen that the confidence interval for 99% confidence level is more than the confidence interval for 95% confidence level.

From the above analysis, we can conclude that with the decreasing of error value and increasing the confidence level, we need large sample size and at that time the confidence interval will include the exact parameter with more confidence level and reliability. On the other hand, taking large sample size is time consuming and costly. So service provider should need to set the value for sample size according to the system condition.

Table 5. CI for location and scale parameters

Parameters	95% confidence interval	99% confidence interval
Location parameter, a	(38.35 , 38.89)	(38.24 , 39.0)
Scale parameter, b	(1.872 , 2.33)	(1.79 , 2.43)

Ⅵ. Conclusion

This paper analyses the most important QoS parameter end-to-end one-way delay for VoIP. We took samples for stationary cases and analyze with different three popular hypothesis tests. In order to estimate the parameter of the logistic distribution, we formulated the equation for the minimum number of sample sizes and confidence interval with error that can be tolerable by the system. This analysis will help to the service provider to know about the end-to-end delay distribution in the real environment and to calculate an appropriate minimum sample size to get an idea about the QoS parameter according to their system limitation.

References

〔1〕 K. Claffy, G. Polyzos, and H.-W. Braun, "Application of sampling methodologies to network traffic characterization," *ACM SIGCOMM Computer Communication Review,* Vol.23, No.4, pp.194-203, 1993.

〔2〕 Xiaoyuan Ta and Guoqiang Mao, “Online End-to-End Quality of Service Monitoring for Service Level Agreement Verification,” *ICON ’06*. Vol.2 pp.1-6.

〔3〕 Wei Zhang, Jingsha He, "Statistical Modeling and Correlation Analysis of End-to-End Delay in Wide Area Network," *IEEE Computer Society*, Vol.3, pp.968-973, August, 2007.

〔4〕 http://www.weibull.com/LifeDataWeb/the_logistic_distribution.htm

〔5〕 Mohd. Noor Islam, Sun Woong Choi, and Yeong Min Jang, “Measurement and Statistical Analysis of QoS Parameters for Mobile WiMAX Network,” *ICACT 2008*, Vol.1, pp. 818-822, February, 2008.

[6] J. D. Bruyne and etc, “Measurement and evaluation of the network performance of a fixed WiMAX system in a suburban environment,” IEEE ISWCS 2008.

[7] B. Sousa and etc, “Experimental evaluation of multimedia services in WiMAX,” Mobimedia 2008.

Mohd. Noor Islam received his B.Sc. in Electrical and Electronic Engineering from Khulna University of Engineering and Technology (KUET), Bangladeshin 2003. Then, he joined as a faculty member in Electrical and Electronic Engineering department of Khulna University of Engineering and Technology (KUET), Bangladesh. He completed his MS from Wireless Networks and Communications Lab. of Kookmin University, Korea in August 2009. His current research interests focus on convergence networks, network resource management, and sensor network.

Yeong Min Jang received the B.E. and M.E. degree in Electronics Engineering from Kyungpook National University, Korea, in 1985 and 1987, respectively. He received the doctoral degree in Computer Science from the University of Massachusetts, USA, in 1999. He worked for ETRI between 1987 and 2000. Since Sept. 2002, he is with the School of Electrical Engineering, Kookmin University, Seoul, Korea. His research interests are IMT-advanced, radio resource management, sensor networks, and convergence networks.