VoWLAN Simulations through Novel Playout Buffer Algorithm

Hua-Ching Chen[1]   Xiao-chao LI[2]   Dong-hui GUO[3]

Department of Electronic Engineering, Xiamen University

Xiamen, Fujian 361005, China

galaxy.km@gmail.com[1], {leexcjeffrey, dhguo}@xmu.edu.cn[2,3]

Abstract. This paper introduces the performance analysis of VoIP applications over WLAN. The factors of VoIP quality include the voice codec, delay, jitter, and packet loss rate. The MOS and E-Model are sequentially considered as subjective and objective assessment of voice quality benchmark to study the relationship between voice quality and end-to-endpoint latency network dynamic environment through the establishment of the mathematical model. This study is also combined with the playout buffer algorithm mechanism to improve the quality of voice communications. In this paper, the researchers proposed a novel play algorithm with pareto distribution mechanism to measure delay, jitter and packet loss rate of VoWLAN voice quality by the MOS scores of different voice codecs. The simulation results proved that the novel adaptive playout buffer mechanism can reduce the latency, jitter, and decrease the packet loss rate and buffer requirements, and correspondently effectively improve the quality of VoIP voice.

Keywords: Simulation, E-model, MOS, Codec, VoIP

1 Introduction

Due to the advances in voice applications services over IP and WLAN technology, it provides low-cost, small-occupied frequency of bandwidth and high efficiency of the network application environment in recent years. Therefore, voice applications over IP service integration in WLAN have got a wide range of response. Because of the dynamic changes of WLAN with respective to many factors, the quality of voice over IP services, including voice codecs, packet loss, the transmission delay and delay jitter cannot be assured. The voice communication is based on UDP transport protocol and each packet of network delay will vary according to the congestion status of WLAN routing. On the transmit side, the speech signal generates packets at regular intervals and sends over the internet to the receiving end. At the receiving end, the voice packets are temporarily stored in a playout buffer prior to playing in order to delay packets playing time, and to suppress the delay jitter. If the setting of playout buffer delay is too long, it can reduce the incidence of packet loss, but it will increase the delay time so as to seriously affect the quality of voice calls. On the other hand, if the setting of playout buffer delay is too short, the too late arriving packets will be not playable and increase the packet loss rate consequently. In order to approach the high-performance requirements of the QoS in voice applications over Wireless LAN (VoWLAN), the playing buffer has to be designed between the playback voice packet delay of receiving end and packet loss rate for enhancing the voice communication service quality. As the wireless network signal is susceptible to environmental interference and power attenuation, the packet transmission delay and packet loss will reduce the voice quality of service. In order to adapt to changes in the dynamic network environment, many researchers have proposed various improved algorithms. For example, Sun [1] proposed the voice quality prediction model based on self-adjust jitter buffer algorithm to adjust the balance delay and reduce packet loss rate voice quality. However, the forecast model ignored the voice quality estimation which affected the performance of these algorithms. Atzori [2] considered the sudden loss patterns and presented self-adjust jitter buffer algorithm.

The application of the E-Model algorithm under voice quality prediction was too complex that limited the time efficiency of the operation. Sanneck [3] used a modified Random Early Detection (RED) algorithm for packet forwarding node, but there was no significant improvement in voice quality. Several real VOIP simulations with NS2 software are developed to verify the real world applications [4-5]. For providing the better solution of the above algorithm problems, the researchers proposed a new Novel Adaptive Playout Algorithm (NAPA). It is based on the mean opinion score (Mean Opinion Score, MOS) of the subjective and objective assessment of voice quality and E-Model. The NAPA combined with the improved VoIP data acquisition mod-

ule and used the Pareto distribution mechanism to provide the high-quality wireless VoIP services through measuring MOS scores of delay, jitter and packet loss rate of VoWLAN voice quality in different voice codecs.

The main structure was divided into two phases in a VoIP model of user activity. The first one is that the beginning of the source encoder and the voice during the conversation activities were based on the choice of codec voice data. The latter one defined the size and frequency of each data packet according to the selected codec, such as G.711, G.723.1, G.729 and GSM.AMR. Such coding rate, frame rate and algorithm delay will cause the occupied different sizes of Bandwidth flow. At the receiving end, VoIP decoder received the voice packets from the network after the properly adjusted parameters setting. The voice packets are broadcasted in the playout buffer for a while to show the playing result. These factors really affected the performance analysis of VoWLAN voice quality by the evolution of MOS model.

In Fig. 1, each typical voice packets possessed with 40-Byte RTP/UDP/IP header and 10 - 30 Byte payload. Such a large head of the table resulted in the occupation of network bandwidth, which greatly affected the voice quality and performance of VoWLAN. Certain domains in these headers were repeated and were a correlation between them, the duplicate information of which could be recovered by the link layer. Thus, if the RObust Header Compression (ROHC) was introduced to compress header for removing redundant information and it could greatly reduce the header's overhead.

IPV4	UDP	RIP	Payload
20B	8B	12B	10~30B
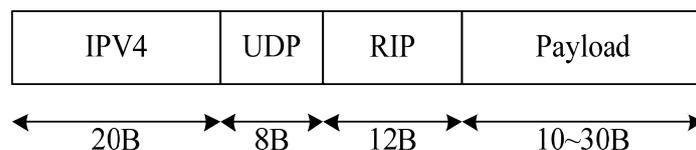
Fig. 1. IPV4 packets of voice

Among which, LH = 40B was the sum of IP/UDP/RTP overhead. Therefore, the efficiency of VoIP voice packet was defined in the IP layer as followings:

$$\eta_{VoIP} = VoIP_{Codec} \frac{(L_{Data} + L_H)}{L_{Data}} \tag{1}$$

The voice signal is the analog waveform, and its voice coding technology is the basis of VoIP.

In order to transfer these analog waveform on the IP address of network, different voice coding could be used to convert voice signal into digital format. According to the different voice coding way, it is occupied by the different bandwidth. In different encoding methods, the transmission of various packet lengths occupied different bandwidth, relatively. Selecting the appropriate encoding algorithm should consider various environmental factors. Additionally, the voice encoder in the voice over the IP services is an important factor of the production of delay variation.

Currently, the applicable VoIP voice encoding are GSM, G.711, G.723.1, G.729A decoder defined by the ITU Remote Communications Organization for Standardization (ITU-T). Table 1 showed the performance comparison between several commonly used speech coding calculus methods [6] [7].

Table 1. Characteristics of Several Voice Codecs

Codec Attributes	GSM	G.711	G.723.1	G.729A
Codec Algorithm	RPE-LPT	PCM	ACELP	CS-ACELP
Bit rate(Kbps)	13.2	64	5.3/6.3	8
Frame interval	20	20	30	10
Payload(Bytes)	33	160	20/24	10
Packets/sec	50	50	33	50
Delay/ms	20	0.75	37.5	15
MOS	3.7	4.1	3.8/3.9	3.9

Speech frame generation was commonly through Voice Activity Detection (VAD) which automatically detected whether the session peak status of both sides was in silent segment or not.

The main function of VAD technology is to automatically and correctly identify the voice and mute segment. In this paper, the researchers simulated Markov chain model of the VAD state during the call. For example, transition methods of four different states including mutual silence, two-way conversation, the average length of stay, and the probability of state transfer. According to the VoIP data acquisition and analysis module [8], the researchers improved the playout buffer adaptive mechanism so that the voice had better performance in the network transmission process this method could be used in a variety of different environments.

The rest of this article was organized as followings. Voice over IP network model was introduced in Section 2. The performance analysis of improving VoIP data acquisition module framework combined with the novel adaptive play algorithm was proposed by the researchers in Section 3. The experimental results are analyzed and explained in Section 4. Finally, the researchers present concluding observations were summarized in Section 5.

2. E-Model Evaluation Methods

Two main factors to affect the VoIP voice quality were voice transmission and network condition. The following were the general performances of the VoIP voice quality: (1) compressed decode (2) Voice Activity Detection (VAD) (3) packet loss rate (4) delay (5) jitter (6) echo. VoIP voice quality measurements were separately divided into subjective and objective evaluation. Subjective evaluation was mainly Mean Opinion Score (MOS) [9]; objective evaluation included Perceptual Speech Quality Measure (PSQM), Perceptual Analysis Measurement, System (PAMS), Perceptual Evaluation of Speech Quality (PESQ) and E-Model [10]. The average subjective value of MOS is widely recognized as the voice quality standards, which were the subjective test results under controlled environmental conditions conducted by many receivers.

E-model is developed by the ITU-T standard [10] [11]. The damaged factors, such as delay, noise, echo, delay jitter, and voice compression, are proposed to analyze the impact of VoIP voice quality. High quality was very suitable for running quality control in the real-time voice communications over IP networks. The final evaluation result of E-Model was expressed by transmission rating factor, R. The R value was a measurement factor of transmission quality which was used to evaluate end-to-end voice quality and to map directly the estimate of the MOS for approaching VoIP call quality. The greater the R value was, the better the voice quality was. Therefore, no matter what methods of measurement must correspond their results to the final average subjective value of MOS. The above methods were indicated as MOS value. The abbreviation and symbol descriptions in E-Model is shown in Table2. Table 3 was the evaluation of the MOS and E-Model for voice subjective and objective views on the overall satisfaction. The voice quality was measured by the average opinion score with five levels to represent the voice quality grades: excellent (4.1 to 4.5 points), good (3.6 to 4.0 points), medium (3.0 to 3.5 points), poor (2.5 to 2.9 points), and bad (2.0 to 2.5 points).

Table 2. Abbreviation and symbol descriptions in E-Model

Abbreviation	description	Abbreviation	descriptions
R_o	Basic signal to noise ratio	I_{olr}	Low OLR values caused by the decline in the quality
I_s	Speech transmission synchronization generated when the damage	I_{st}	Damage which is caused by the non-appropriate telephone noise
I_d	Speech signal delay caused by the damage	I_q	Injury due to the distorted quantification
$I_{e\text{-}eff}$	Damage to the low bitrate codecs	I_{dte}	Assessment of losses caused by sender's echo
A	Benefit coefficient	I_{dle}	Damage caused by the echo of the receiving end
SLR	Send loudness levels	I_{dd}	Damage due to long absolute delay
RLR	Receive loudness levels	Ie	Equipment damage coefficient
$STMR$	Sidetone masking level	Bpl	Packet loss, strength coefficient
$LSTR$	Sidetone level	Ppl	Random packet loss probability
Ds	Send side telephone D value	$BurstR$	Sudden ratio
Dr	Receiving side telephone D value	qdu	Quantization distortion unit quantity
$TELR$	The speakers echo intensity level	Ps	Sending side indoor noise
T	The average one-way delay of the echo channel	Pr	Receiving side indoor noise
Ta	Anechoic connection delayed	I_r	Damage factors which meant coding injury without packet loss

Table 3. Equivalent R values into estimate MOS

Level	R-value(lower limit)	MOS(lower limit)
Excellent	90~100	4.1~4.5
Good	80~90	3.6~4.0
Medium	70~79	3.0~3.5
Poor	60~69	2.5~2.9
Bad	50~59	2.0~2.5

Voice services in the analytical model of the WLAN system applications mainly adopted the two-state Markov ON/OFF model of the Poisson process. In order to analyze the WLAN voice quality from endpoint to endpoint, ITU-T G.107 proposed a kind of the E-Model analysis tools [10] [11] which could predict the performance analysis of voice quality through network parameters calculation and some other system parameters. The main logically assessment methods were shown in Fig. 2. The user's silent time and the status probability were

affected by end-to-end delay, jitter, packet loss and other factors by Hammer's analysis[12]. In a two-way dialogue session, the streams were independent of each other. Dialogue between two sides could be one of these four statuses, including mutually silent, the two-way communication, an average length of stay, and the state transition probability calculation.
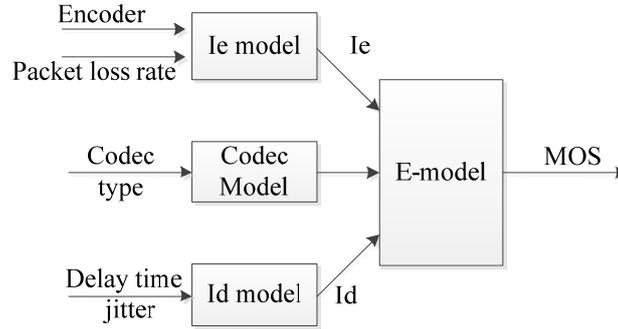
Fig. 2. E-Model logic diagrams

The E-Model's voice quality definition was determined by the voice transmission parameters whose quality loss was formula a parameter set to calculate the delay, packet loss, equipment impairment factors. The result of E-Model calculations was the transmission performance coefficient level [13], presented as R factor.

$$R = R_o - I_s - I_d - I_e + A \tag{2}$$

In Eq (2), Ro represented the basic signal-noise ratio includes the loudness evaluation value (*SLR*, *RLR*), background noise (P_s, P_r), telephone D value (D_s, D_r), etc. I_s showed the influence of synchronization injury, quantification, connection and echo caused by voice transmission. I_d said the damage caused by the delay of the voice signals transmission, including call echo and the impact of interactive loss. $I_{e\text{-}eff}$ offered effective equipment damage coefficient which was the damage made by low bit rate codec, including damage of the randomly distributed packet loss. The symbol "A" presented the advanced profit coefficient which was the compensation of damage coefficient. The calculation formula of I_d was listed as bellows:

$$I_d = 0.024d + 0.11(d - 177.3)H(d - 177.3) \tag{3}$$

Among this, d was the network one-way delay, i.e. the playout delay, which was the sum of network delay and playout buffer delay. $I_{e\text{-}eff}$ presented the damage factor which is caused by the device, such as the factors of injury and loss of the encoder. Its formula was listed as below:

$$I_{e-eff} = I_e + (95 - I_e)\frac{Ppl}{\dfrac{Ppl}{BurstR} + Bpl} \tag{4}$$

P_{pl} was the percentage of packet loss rate. I_r was the equipment damage factors which meant coding injury without packet loss. B_{pl} is considered as the packet loss robustness factor. Zlatokrilov [14] thought that, in VoIP applications, the burst packet loss is widespread phenomenon. So the adopted -Model to predict the voice quality had to consider the influence of BurstR parameters. Two-state Markov model was the packet loss process modeling by Raake [15]. BurstR could be calculated by Eq (5):

$$BurstR = \frac{1}{p+q} = \frac{Ppl/100}{p} = \frac{1 - Ppl/100}{q} \tag{5}$$

Among this, p states the transition probability from the receiving to losing status in the Markov model. On the contrary, q meant the transition probability of the status from loss to receiving. The calculation of I_s is by formulas,

$$Is = Iolr + Ist + Iq \tag{6}$$

Where I_{olr} is the quality declination caused by the low loudness and evaluation value. I_{st} means the damage which is caused by the non-appropriate telephone noise. I_q presented the injury due to the distorted quantification.

The parameters I_s are affected by the factor of loudness the evaluation value (*SLR*, *RLR*), background noise (P_s, P_r), telephone D value (D_s, D_r), speaker's echo sound rating (*TELR*), one-way delay time (*T*), side-tone mask rating (*STMR*), and distorted quantification degree (*qdu*). Id showed the delay damage which is caused by the speech signal. There are further divided into three factors, I_{dte}, I_{dle} and I_{dd}. The calculation of I_d is

$$Id = Idte + Idle + Idd \tag{7}$$

Among this, I_{dte} expressed the assessment of losses caused by sender's echo. I_{dle} said the damage caused by the echo of the receiving end. I_{dd} presented the damage due to long absolute delay. When the echo was cancelled, the absolute delay was generated. The parameters I_d were affected by the one-way delay (T), the absolute delay (T_a), the speaker sound rating (*TELR*), side-tone mask rating (*STMR*), loudness evaluation value (*SLR*, *RLR*), background noise (P_s, P_r), and telephone D value (D_s, D_r). I_e is the damage factor caused by the low bit rate coding and packet loss. The MOS scores are converted through the range of parameters R.

$$MOS = \begin{cases} 1 & R < 0 \\ 1 + 0.035R + R(R-60) \times 7 \times 10^{-6} & 0 < R < 100 \\ 4.5 & R > 100 \end{cases} \tag{8}$$

Due to the main effects of voice quality by the variant network states, the proposed model calculation of the default parameters in our research is simplified by G.107[16].

$$R = 93.2 - Id - Ie - Ij + A \tag{9}$$

2.1 Model analysis

Based on the binary exponential of backoff model, Bianchi [17] proposed the Markov model in the saturated case. The main assumptions was that the probability of data sent by each site as a given constant. The probability of conflict in the transmitted data packets are also a given constant. There are nothings to do with the number of the packet retransmission. The Markov state transition diagram was offered in Fig. 3.
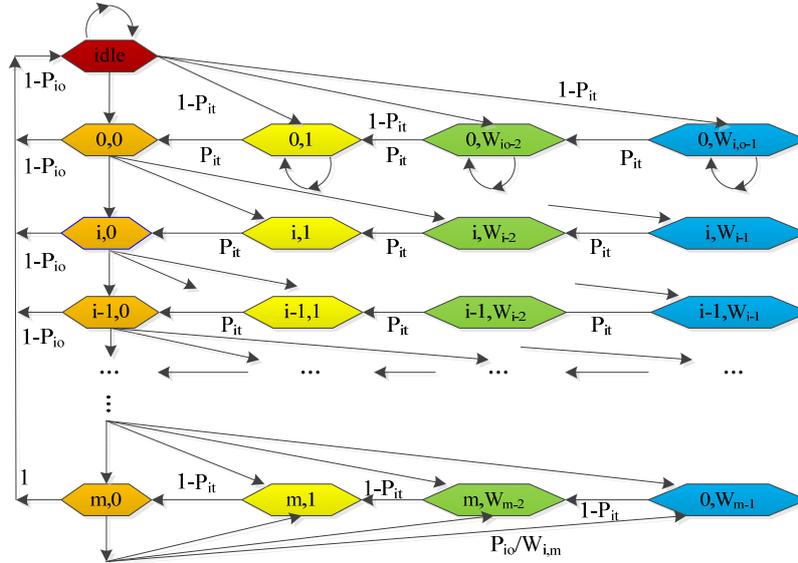
Fig. 3. Transition diagram of the proposed Markov chain model for AC[i].

Assumed that p was the conditional probability of collision whose value was constant and independent of each other. The collision probability was not related to the number of collisions in each transmission. Then, the backoff algorithm of the non-Markov processes could transform into a two-dimensional discrete-time Markov $\{s(t), b(t)\}$. Based on the illustration of Fig. 3, Markov's analysis of state transition function was listed below:

$$\begin{cases} p\{i,k \mid i,k+1\} = 1 & k \in (0, W_i - 2) \quad i \in (0, m) \\ p\{i,k \mid i,0\} = (1-p)/W_0 & k \in (0, W_0 - 1) \quad i \in (0, m) \\ P\{i,k \mid i-1,0\} = p/W_i & k \in (0, W_i - 1) \quad i \in (1, m) \\ P\{m,k \mid m,0\} = p/W_m & k \in (0, W_m - 1) \end{cases} \tag{10}$$

The researchers used a simplified expression as followings:

$$p\{i,k \mid i_0,k_0\} = p\{s(t+1) = i_1, b(t+1) = k_1 \mid s(t) = i_0, b(t) = k_0\} \tag{11}$$

In Eq (10), the first equation represented a probability of the backoff timer minus 1 was 1 at the moment of the beginning of each time slot. The second equation showed the new frame retracted degree began with 0 after successfully transmitting a frame. The counter was randomly selected an integer value from (0,W_{0-1}) as the re-

traction time. The third equation expressed that after an unsuccessful transmission, the backoff counter added one level on the retracted series. The backoff value selected a random integer value from $(0, W_{0-1})$; the last equation was the selected method from the value of $(0, W_{m-1})$ in an unsuccessful transmission when the retracted order had reached its maximum.

After a long observation for the dimensional random process of site backoff time, Markov chain showed a steady-state distribution which listed as below: $b_{i,k} = \lim_{\tau \to \infty} P\{s(t) = i, b(t) = k\}$, $i \in (0, m), k \in (0, W_i)$. It Stated that $\{i, k\}$ was the stability probability, and $b_{i,k}$ meant the probability of each state in the Markov chain. According to Markov state transition diagram, the result would be:

$$\begin{cases} b_{i-1,0} \times p = b_{i,0} \to b_{i,0} = p_i b_{0,0} & 0 < i < m \\ b_{m-1,0} \times p = (1-p) \times b_{m,0} \to b_{m,0} = \dfrac{p^m}{1-p} \times b_{0,0} \end{cases} \tag{12}$$

That is,:

$$\begin{cases} b_{i,0} = p_i b_{0,0} & 0 < i < m \\ b_{m,0} = \dfrac{p^m}{1-p} \times b_{0,0} \end{cases} \tag{13}$$

Based on the regularity of the Markov chain, a state point $b_{i,k}$ expression could be written as followings:

$$b_{j,k} = \frac{W_i - k}{W_i} \times \begin{cases} (1-p) \sum_{j=0}^{m} b_{j,0} & i = 0 \\ p \times b_{i-1,0} & 0 < i < m \\ p \times (b_{m-1,0} + b_{m,0}) & i = m \end{cases} \tag{14}$$

According to Eq (13), $\sum_{i=0}^{m} b_{i,0} = b_{0,0} / (1-p)$, and then Eq (14) formula expression was the followings:

$$b_{i,k} = \begin{cases} \dfrac{W_i - k}{W_i} b_{i,0} & i \in (0, m) \quad k \in [0, W_i - 1] \end{cases} \tag{15}$$

The steady-state distribution possessed with the normalization condition. The sum of all state probability is one. By Eq (13) and Eq (15), the researchers got $\sum_{i=0}^{m} \sum_{k=0}^{W_i-1} b_{i,k} = 1$, and the $b_{i,k}$ is presented by

$$1 = \sum_{i=0}^{m} b_{i,0} \sum_{k=0}^{W_i-1} \frac{W_i - k}{W_i} = \sum_{i=0}^{m} b_{i,0} \times \frac{W_i + 1}{2} = \frac{b_{0,0}}{2} \left[CW_{min} \left(\sum_{i=0}^{m-1} (2p)^i + \frac{(2p)^m}{1-p} \right) + \frac{1}{1-p} \right] \tag{16}$$

$$b_{0,0} = \frac{2(1-p)}{(CW_{min} + 1) + pCW_{min}(2p)^{m-1}} \tag{17}$$

The η is defined as the transmission probability of a transmission packet in any randomly selected time slot. Without the consideration of the retracted order count value, the data packet began to transmit when backoff time of the backoff calculator was reduced to 0.

$$\eta = \sum_{i=0}^{m} b_{i,0} = \frac{b_{0,0}}{1-p} = \frac{2}{pCW_{min}(2p)^{m-2}} \tag{18}$$

When $m = 0$, there was no exponential backoff. The p was the probability of η results. Eq (18) could be simplified as belows:

$$\eta = \frac{2}{CW_{min} + 1} \tag{19}$$

From Eq (18), it could be noted that η is depended on the conditions of collision probability p in Eq (19). The probability of p was the collision probability of data packets transmission. In one slot, at least one of n-1 of the remaining station is transmitted. The above assumption was under the same stable condition. In the steady state, the conflict probability p could be expressed when sending packets at any point.

$$p = 1 - (1 - \eta)^{n-1} \tag{20}$$

η and p are the two unknowns' values for the Eq (18) and Eq (20) in the nonlinear system. Based on the reversed of Eq (20), the researchers get $\eta * (p) = 1 - (1-p) \, 1 \, / \, (n-1)$ which was a continuous and monotonically increasing function when the range of $p \in (0,1)$. From the definition of equation of η $(p)\eta^*(0)=0$ and $\eta^*(1)=1$, it is satisfying to the continuous range of $p \in (0,1)$. The continuity of the critical value is selected as p = 1/2. Eq (18) could be presented as Eq (21):

$$\eta(p) = \frac{2}{1 + CW_{min} + pCW_{min} \sum_{i=0}^{m-1} (2p)^i} \tag{21}$$

When $p = \dfrac{1}{2}$, $\eta(\dfrac{1}{2}) = \dfrac{2}{1 + CW_{min} + (pCW_{min} / 2)}$, $\eta(p)$ was a continuous function, and a monotonically decreasing function. The definition of $\eta(0) = \dfrac{2}{CW_{min} + 1}$, $\eta(1) = \dfrac{2}{1 + 2^m CW_{min}}$. The backoff algorithm was not adaptable and independent of η and p When $m = 0$. So that $\eta = \dfrac{2}{CW_{min} + 1}$.

2.2 Delay Analysis

Due to the dynamic behavior of the WLAN network, it contained the dynamic process of change during voice conversations. The voice conversations distribution states were related to its various network environments. For example, in a certain time period, distribution A could be fitted well, but, in another time period, distribution B or C could be a better fit. It was difficult to describe accurately the distribution of voice conversations in every environment. The current consensus was that the node of voice conversations generally subjected to heavy tail distribution. The most common delay distributions were exponential distribution [18], the Weibull distribution [19] and Pareto distribution [20]. It is noted that the Pareto distribution for VoWLAN traffic delay distributions approach the lower complexity and the suitable regulation [21] [22] [23].

At the MAC layer, media access time delay was the summatio
n of time period when the data frame was successfully sent out and competed for Letter Road in a node. If the RTS/CTS access method were adopted, it also included the time for the successful exchange of the RTS/CTS way. Its calculation is as follows:

$$E_M = E_{avg} \times E_{slot} \tag{22}$$

E_M represented the average system time slot number that nodes had to wait for in the completion of a successful transmission. E_{slot} said the average duration of a system slot, which could be expressed below:

$$E_{slot} = (1 - P_{tr})\sigma + P_{tr} P_s T_s + P_{tr}(1 - P_s)T_c \tag{23}$$

2.3 Packet loss Analysis

Packet Loss meant that the voice packets were discarded because of unsuccessful delivery to the receiving end on time.

In E-Model approach, the percentage of packet loss rate (P_{pl}) was the total packet loss rate. It includes packet loss rate Ppl_n which is caused by network traffic and packet loss rate P_{pln} which is affected by the jitter buffer. The calculation of packet loss rate is

$$Ppl = 100 \times (Ppl_n + Ppl_b) \tag{24}$$

Among this, P_{pln} could be calculated directly in the process of packet transmission. In order to obtain the jitter buffer packet loss rate P_{plb}, the delay d had to develop the right distribution model. Therefore, the P_{plb} could be calculated by the delay Cumulative Distribution Function (CDF):

$$Ppl_b = P(X \geq d) = 1 - F(d) \tag{25}$$

$F(x) = P(X \leq x)$ is the cumulative distribution function of the delay distribution. Pareto distribution of the cumulative distribution function could be expressed by following formula:

$$F(x) = 1 - \left(\frac{\kappa}{x}\right)^{\alpha}, x \geq \kappa \tag{26}$$

The packet loss rate P_{plb} could be calculated by the affection of jitter buffer as following:

$$Ppl_b = 1 - F(d) = \left(\frac{\kappa}{d}\right)^{\alpha} \tag{27}$$

Pareto distributed parameters κ and α could be estimated by the number of ω which had received the packet network delay data. The parameters are estimated as following:

$$\kappa = \min_{j \in w}\{n_i\} \tag{28}$$

$$\alpha = \omega \left[\sum_{i=1}^{w} \log\left(\frac{n_i}{\kappa}\right)\right]^{-1} \tag{29}$$

n_i means the i-th received packet network delay. Based on the Delay distribution model, the total percentage of the packet loss rate P_{pl} could be expressed as a function of playback delay d.

3   E-Model basis novel adaptive playout buffer mechanism

3.1   Novel adaptive voice playout buffer algorithm

For the true reflection of the voice quality evaluation in the wireless LAN, the improved evaluation model was proposed according to analyze the factors of VoIP service quality. The original voice data network evaluation was also adopted to combine the different playout buffer algorithms for getting the average opinion score of the MOS voice service in wireless local area network. This paper presented a Novel Adaptive Playout Algorithm (NAPA) to provide the high-quality wireless VoIP services. Based on the data of MOS subjective and objective assessment of voice quality conditions and the E-Model, the Playout Buffer calculus law mechanisms was combined with to achieve the high-quality voice communications. The following steps were the performance assessment for the endpoint to the endpoint analysis.

There is causal relationship existed among delay, packet loss rate, and jitter buffer size. The objective of balancing the adaptive jitter buffer size is to minimize the length of the delay and packet loss number under the condition of jitter elimination. The NAPA algorithm is to ensure the quality of voice transmission by reducing the number of voice packet loss in the maximum extent possible within the reasonable limits. The latency and packet loss are controlled under the 400ms and 10%, which is still within the acceptable range for voice quality. In the other hand, the performance could not meet the requirements while the voice delay time is over 400ms and packet loss rate over 10%.

The algorithm's concept of flowchart was shown in Fig. 4. The maximum acceptable packet loss rate is 10% and range of the acceptable voice delay Time (0, 400). It is clear that the premise of the entire algorithm must ensure that the delay within the 400ms reasonable range. In other words, the proposed algorithm increases the size of the buffer B_{jitter} to keep the current time delay $P_{lossrate}$ <10% when loss rate $P_{lossrate}$ met $P_{lossrate}$ > 10% at the condition of delay was less than 400ms. The $P_{lossrate}$ could be properly over 10% to fit the maximum 400ms time delay when meets the condition of $P_{lossrate}$ <10% and time delay over 400ms. The B_{jitter} can be decreased until the time delay was within the reasonable range when time delay over 400ms. If packet loss rate and delay time are still not acceptable, the level of voice MOS is less 3 and package voice will be lost in this time. From the design of flow chart, the self-adapt algorithm would be adjusted per sending voice packets until the end of the call.
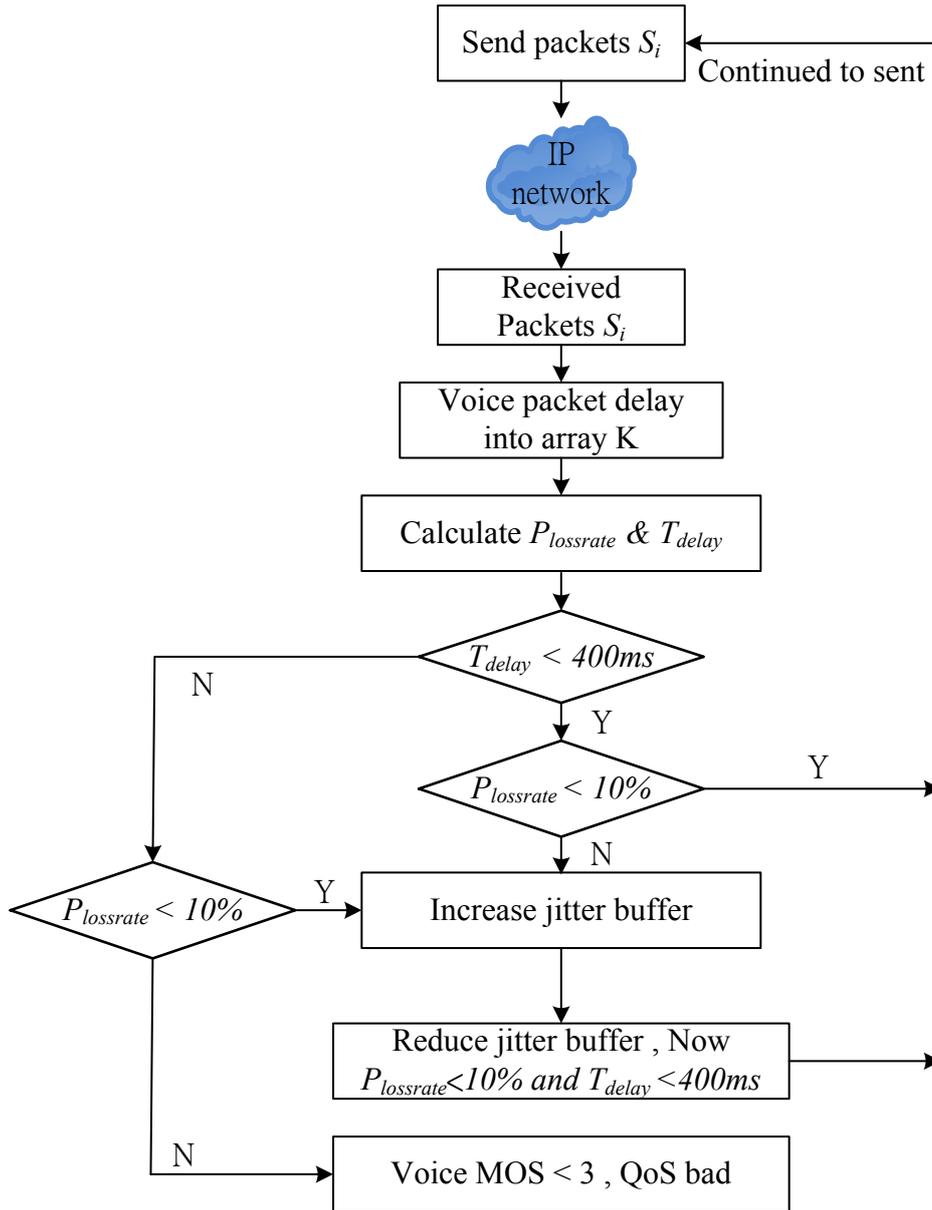
Fig. 4. NAPA algorithm flow chart

3.2 Novel Adjustment Stages

The direct correlation was existed between jitter buffer size and playback delay. In order to obtain a balance between time delay and packet loss, the goals of NAPA algorithm was to dynamically adjust the size of the jitter buffer, to set a reasonable time delay d for adapting the changes in network environment, and to optimize the voice quality. Each voice call was consisted of voice segment and mute segment, and each voice segment was called a voice talkspurt. Mute segment was existed between two talkspurts. Therefore, in order to ensure the immediacy of jitter process, the NAPA algorithm would calculate at the beginning of each talkspurt for adaptively adjusting to achieve the best delay. The process was listed as followings:

(1) According to the historical delay information of the latest m data packets, it was estimated the parameters κ and α of Pareto delay distribution;

(2) Calculated the packet loss rate Ppl_b caused by jitter buffer and packet loss rate Ppl_n caused by network traffic;

(3) Using E-Model, the functional relationship $f_{MOS}(d)$ could be extracted on the I_{e-eff} and I_d basis.

(4) search for the optimal playback delay $d_{optimal}$ for the current talkspurt:

$$d_{optimal} = d_{\arg(\max)} \times f_{MOS}(d) \qquad (30)$$

The voice quality prediction and adaptive adjustment were the two key issues of the algorithms in this article. Among them, the voice quality measurements needed to establish the delay distribution model to estimate the packet loss rate, and adopted E-Model to calculate the voice quality under sudden packet loss mode. Delay distribution model must comply with the characteristics of the network delay and possess with lower complexity, whose parameter estimation must gain a balance between accuracy and sensitivity. Hence, a reasonable size had to be set when using historical data window of parameter estimation. If the data window was too big, the sensitivity of delay distribution model would be low and not enough to adapt to the delay variation; if the data window was too small, it would be difficult to accurately estimate the packet loss rate.

Setp 1)   When $d \le 100ms$, the parameters I_d value was very small, almost no delay damage. So the play-back delay d should be set to greater than 100 ms in order to prevent the loss damage;

Setp 2)   According to VoIP systems the voice coding frame length fr, the distance of the search calculation was set. Calculated $d_m = 100 + m * fr (m = 1, 2, ..., n)$, the corresponding function values $f_{MOS}(d_m)$. Took $f_{MOS}(d_{optimal}) = \max\{f_{MOS}(d_m) | m = 1, 2, ..., n\}$ to obtain the optimal playback delay $d_{optimal}$;

Setp 3)   User perception of voice quality had been poor when $d \ge 400ms$. Therefore, the playout delay should be set to within 400 ms. That was $n = \left| \dfrac{400 - 100}{fr} \right|$. Through adaptive adjustment of the above, the highest predictive value of voice quality was selected as the final result.

4   Simulation performance assessment

In order to assess the performance of NAPA, the network simulator NS2, (Network Simulator 2) [24] software was used to simulate several experiments in this paper. The NAPA algorithm compared with Atzori algorithm, optimal playback algorithms were examined in these experiments. The acquired VoIP data module is referenced by Bacioccola in University of Pisa in Italy from NS2 environments [4] [25] in this experiment. The simulated VoIP network is disposed on 802.11b wireless LAN whose data transfer rate is 11 Mbps.

The utility of NS2 simulator is proposed to measure the characteristics of the VoIP channels. Such performances are delay, jitter, packet loss rate, throughput analysis in the VAD model analysis of various nodes with different codecs (GSM AMR, G.711, G.723.1, G.729A). The delay of voice traffic and MOS [26] [27] were fully analyzed to evaluate performance between the internal network nodes. This experiment is shown that 50 random mobile nodes are deployed in a 500m * 500m area in the Ad Hoc network topology. The source node sent voice data flow with a constant bit rate (CBR) to destination node i. These signals could achieve the communication between each other by sending the voice data through the middle of the nodes 1,2, ... Simulation time is selected as 100 seconds. The experiment of each case was repeated 20 times to gain the authenticity of the experimental data.

Experiment 1 Delay evaluation and analysis
In order to understand the delay changes of voice transmission in a WLAN environment, the researchers changed the network transmission time and analyzed the MOS value of voice transmission in different time delay with different codecs. The researchers also used the incremental method to test the time delay of speech transmission, and analyzed the voice services quality of MOS value under different time delay changes. The voice service quality performance of time delay variation were analyzed and compared by three playout buffer algorithms in WLAN, including NAPA, Optimal, and Atzori. As the voice packet delay increased from 10ms to 100ms, the comparative analysis of the MOS values of the delay variation in different codecs would be done to understand the performance of voice service quality of each encoder.

GSM.AMR   G.711   G.723.1   G.729A
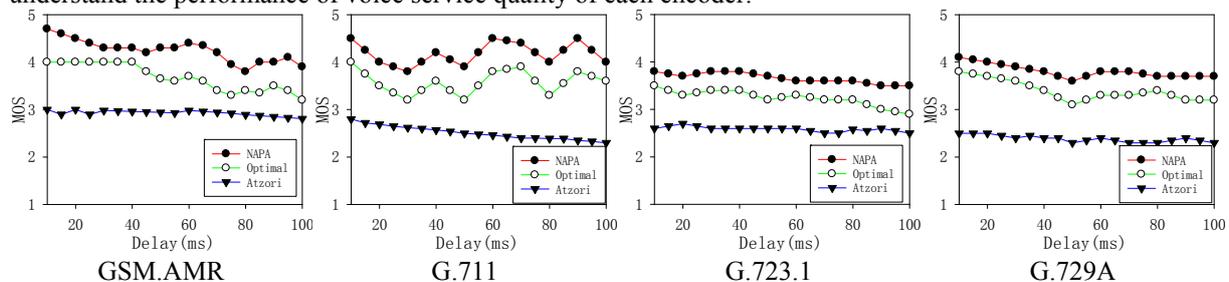
Fig. 5. Average Delay vs. MOS

Table 4. MOS value of the definition and Average Transmission Delay

Algorithm	GSM.AMR	G.711	G.723.1	G.729A
NAPA	4.27	4.38	3.96	4.01
Optimal	3.83	3.70	3.45	3.72
Atzori	2.95	2.45	2.59	2.45

In Fig. 5, it was known that different time delays had different influence on curve change of each encoder. It also showed that the curve change of the frame transmission in GSM.AMR, G.711, G.723.1, G.729A when applying NAPA. The MOS values of different speech codec were relatively higher and suitable for using in different codecs. But the MOS value applied Atzori algorithm was significantly lower. The time delay change of four codecs/decoders and their curve performance of MOS value were significantly affected by the increase of time delay, and the voice quality was pretty unacceptable that was not suitable for WLAN different delay changes in transmission. In addition, from Table 4, the voice service quality of different delay variation was very clear. NAPA still maintained a good quality of voice calls as the increase in the amount of WLAN delay variation. In overall analysis of delay evaluation, NAPA had better performance than the other two and were suitable for the VoWLAN application in different codecs.

Experiment 2, Evaluation and analysis of error rate

In order to understand the change of the WLAN voice transmission error rate, this paper obtains the analyzed MOS in the individual packet loss rate of different codes based on the random variable method. At first, the voice service quality performance of three compared playout buffer algorithms include NAPA, Optimal and Atzori, are applied in WLAN. As the voice packet error rate increasing from 1% to 10%, the change of error rate in the different codecs were compared and analyzed to get its MOS score values for understanding the performance of each encoder's voice service quality.
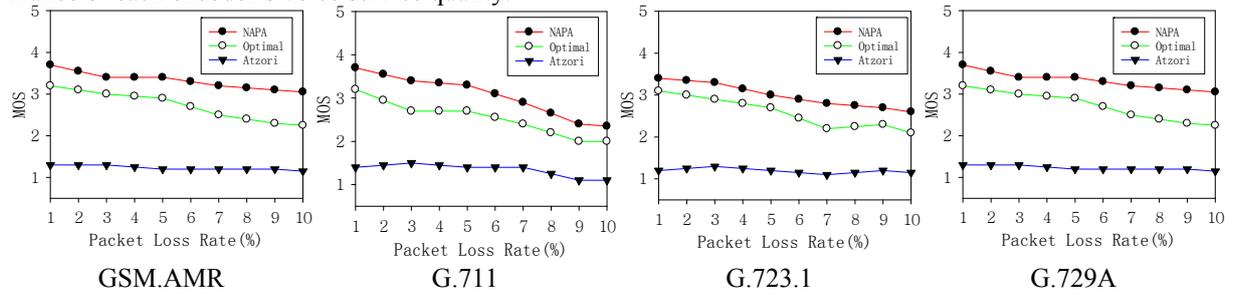
Fig. 6. Average Packet Loss Rate (PLR)

Table 5. MOS value of the definition and Average Packet Loss Rate (PLR)

Algorithm	GSM.AMR	G.711	G.723.1	G.729A
NAPA	3.36	3.20	3.26	3.24
Optimal	2.96	2.95	2.89	2.92
Atzori	1.41	1.39	1.36	1.25

Fig. 6 illustrates that these different error rates had its respective changes of encoders' curve. The NAPA obtained a relatively higher MOS value from the curve changes of packet loss rate in GSM.AMR, G.711, G.723.1 G.729A when applying in each codecs. This meant that it was suitable for different codecs. However, when applying Atzori and Optimal algorithm, the MOS value was significantly lower. The four codecs/decoders' loss rate and their performance curve of MOS value were significantly affected by the network package loss rate, and their voice qualities were serious unacceptable. Further, when the network packet loss rate reached 10%, the MOS value of NAPA was still at 3 or more, and its voice quality services were in accordance with ITU-T standards. But the MOS value of Atzori algorithm were only about 2.2, and the voice data were not appropriate to transmit under the highly changeable error rate. In Table 5, it was clear that the voice call quality had different performance because of the lack of adaptive playout buffer. Additionally, the network delay would also reflected in the MOS value which was the representative of the voice calls quality.

5 Conclusions

Influenced by factors such as the voice codec, delay, jitter, and packet loss rate. Therefore, the researchers used the subjective and objective assessment condition of voice quality in MOS and E-Model as the basis, established a mathematical model, and then combined with the playout buffer algorithm mechanism to measure the quality of voice communications. Additionally, the researchers also offered an idea of combining a novel adaptive play-

back algorithm with improved VoIP data acquisition module in order to obtain the MOS value of VoWLAN's voice quality, delay, jitter, and packet loss rate in different voice codecs.

The simulated experiment showed that VoWLAN voice quality performance by the novel adaptive playout buffer algorithm obtains the better performance in different codecs example. The developed algorithms can be able to reduce the delay, jitter, packet loss rate and buffer requirements. Hence, in overall, it could promote the voice quality of VoIP and gain user's recognition in VoWLAN applications.

Some interesting topics in the optimal combinations research and evaluate the effect of different performance analysis in packet loss rate, codec type and delay time to solve the overall integration problem.

Acknowledgements

References

[1] L. Sun, C.E. Ifeachor, "Voice Quality Prediction Models and Their Application in VoIP Networks," *IEEE Transactions on Multimedia*, Vol. 8, no. 4, pp. 809-820, 2006.

[2] L. Atzori, M. L. Lobina, M. Corona, "Playout Buffering of Speech Packets Based on A Quality Maximization Approach," *IEEE Transactions on Multimedia*, Vol. 8, No. 2, pp. 420-426, 2006.

[3] H. Sanneck, N. T. L. Le, A. Wolisz, G. Carle, "Intra-flow Loss Recovery and Control for VoIP," in *Proceedings of 9th ACM International Conference on Multimedia*, pp. 441-454, 2001.

[4] A. Bacioccola, C. Cicconetti, G. Stea, "User-level Performance Evaluation of VoIP Using ns-2," in *Proceedings of 2nd International Conference on Performance Evaluation Methodologies and Tools (ValueTools'07)*, pp. 11-20, ACM Press, 2007.

[5] M. M. Andreozzi, D. Migliorini, G. Stea, C. Vallati, "Ns2Voip++, An Enhanced Module for VoIP Simulations," in *Proceedings of 3rd International ICST Conference on Simulation Tools and Techniques*, pp. 48-50, ACM Press, 2010.

[6] A. Mohd, O. L. Loon, "Performance of Voice over IP (VoIP) over a wireless LAN (WLAN) for Different Audio Voice Codecs," *Jurnal Teknologi*, Vol. 47, pp. 39-60, 2007.

[7] S. M. Oh, J. H. Kim, "Application-Aware Design to Enhance System Efficiency for VoIP Services in BWA Networks," *IEEE Transactions on Multimedia*, Vol. 13, No. 1, 2011.

[8] http://cng1.iet.unipi.it/wiki/index.php/Ns2voip

[9] ITU- T Recommendation, *P.862: PERceptual Evaluation of speech quality (PESQ), An Objective Method for End-to-end Speech Quality Assessment of Narrowband Telephone Networks and Speech Codecs*, 2001.

[10] ITU- T Recommendation G. 107. *The E-Model, A Computational Model for Use in Transmission Planning*, 2000.

[11] ITU-T Recommendation G.108, *Application of the E-model: A planning Guide*, Sep 1998.

[12] F. Hammer, P. Reichl, A. Raake, "The Well Ttempered Conversation: Interactivity, Delay and Perceptual VoIP Quality," in *Proceedings of IEEE International Conference on Communications*, pp. 244-249, IEEE Press, 2005.

[13] V. A. Reguera, F. F. Á. Paliza, W. Godoy Jr., M. Evelio, G. Fernández, "On the Impact of Active Queue Management on VoIP Quality of Service," *Computer Communications*, Vol. 31, pp. 73-87, 2008.

[14] H. Zlatokrilov, H. Levy, "Packet Dispersion and the Quality of Voice over IP Applications in IP Networks," in *Proceedings of 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004)*, pp. 1170-1180, IEEE Press, 2004.

[15] A. Raake, "Short-and Long-term Packet Loss Behavior: Towards Speech Quality Prediction for Arbitrary Loss Distributions," *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 14, No. 6, pp. 1957-1968, 2006.

[16] ITU-T Recommendation G.107, *The E-Model, A Computational Model for Use in Transmission Planning*, 2005.

[17] G. Banchi, "Performance Analysis of the IEEE 802. 11 Distributed Coordination Function," *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 3, pp. 535-547, 2000.

[18] P. T. Brady, "A Model for Generating on-off Speech Patterns in Two-way Conversation," *Bell System Technical Journal*, Vol. 48, pp. 2445-2472, 1969.

[19] D. Stutzback, R. Rejaie, "Understanding Churn in Peer to Peer Networks," in *Proceedings of 6th ACM SIGCOMM on Internet Measurement Conference (IMC'06)*, pp. 189-202, ACM Press, 2006.

[20] C. N. Chuah, R. H. Katz, "Characterizing Packet Audio Streams from Internet Multimedia Applications," in *Proceedings of IEEE International Conference on Communications (ICC 2002)*, pp. 1199-2203, IEEE Press, 2002.

[21] Z. P Fu, H. M. Wang, D. X. Shi, P. Zou, "A Review on the Churn Character of P2P Networks," *Chinese Journal of Computers*, Vol.9, pp. 1563-1577, 2011.

[22] W. Zhang, J. He, "Modeling End-to-end Delay Using Pareto Distribution," in *Proceedings of IEEE Second International Conference on Internet Monitoring and Protection* (*ICIMP 2007*), pp. 21-24, IEEE Press, 2007.

[23] T. Cai, Z. Long, X. WU, "Adaptive Jitter Buffer Algorithm of VoIP Based on Speech Quality Prediction," *Computer Engineering and Applications*, Vol. 47, No. 10, pp. 63-66, 2011.

[24] http://www.isi.edu/nsnam/ns

[25] http://info.iet.unipi.it/~cng/ns2voip/

[26] S. Jelassi, G. Rubino, H. Melvin, H. Youssef, G. Pujolle, "Quality of Experience of VoIP Service: A Survey of Assessment Approaches and Open Issues," *IEEE Communications Surveys & Tutorials*, Vol. 14, No. 2, pp. 491-513, 2012.

[27] Y. Lee, S. G. Yoo, S Lee, "An Efficient Detection and Management of False Accusation Attacks in Hierarchical Ad-Hoc Networks," *KSII Transactions on Internet and Information Systems*, Vol. 6, No. 7, pp. 1874-1893, 2012.