The Study of Internet Telephony based on Real Time Protocol

以即時協定為基礎的網際網路電話研究[1]

Liang-Fu You
System Simulation Office, System Development Center
Chung-Shan Institute of Science & Technology
david@royals.ee.nctu.edu.tw

Abstract

The computer telephony gateway (or called internet telephony gateway) is the bridge between telecommunication and internet. Usually using a computer added a telephony interface card becomes a computer telephony gateway. The paper first offers two architecture of the internet telephony gateway using modems, a sound card and a network interface card that every personal computer already has. With the internet telephony software "SpeakFreely" we experimented and found out that the packet loss took place in insufficient bandwidth and the spike took place in network jam. The paper issued an algorithm for detection and solution of the spike or bandwidth insufficiency. With such algorithm to improve the voice quality of internet telephony.

Keywords:

Real time protocol (RTP), Computer telephony gateway, Hybrid transformer, Adaptive playout buffer, Spike

摘　要

電腦電話閘道器(或稱為網際網路電話閘道器)是電信與網際網路的溝通橋樑，通常使用電腦加裝電話語音卡，成為電腦電話閘道器。本論文首先提出兩種利用每部個人電腦通常就具備的數據機、聲霸卡和網路卡，來完成電腦電話閘道器的功能架構。經由網際網路電話軟體SpeakFreely之實測結果可知，語音封包的遺失發生於頻寬不足，極長的語音延遲(spike)發生於網路擁塞。本論文提出由即時協定的語音封包接收情形，計算出目前之網路頻寬的數學公式，以便選用最適合的語音壓縮技術；並提出改善的spike偵側公式，於spike發生的最初期，立即改用較高壓縮率的語音壓縮方法。藉此spike和頻寬不足之偵測及解決的演算法，來改善網際網路電話的語音品質。

關鍵詞：

即時協定, 電腦電話閘道器, 混合式變壓轉換器, 動態撥放暫存，極長的語音延遲

[1] 本文獲中華民國資訊學會2001年最佳碩博士論文獎-碩士論文佳作

1.Introduction

Now the internet is a part of the telecommunication, but in the future telecommunication will be a part of the internet. Traditionally we used telephone to communicate with each other. The Public Service Telephone Network (PSTN) provided switching system to form the telecommunication system. After the invention of computer, we want the jobs of complex calculation to be done by computer and the PSTN still provide switching function. So the technology of Computer Telephony Integration (CTI) occurs and many CTI applications are applied, such as call center, interactive voice response, internet telephony, etc.

The internet telephony, also called VoIP, IP Phone or IP telephony, can cost down the telephone fees, but the voice quality are not as good as PSTN. Many people study new methods to improve the voice quality in internet telephony, like noise reduction, echo canceling, voice compression, adaptive playout buffer, Real Time Protocol (RTP), and so on.

There are five types of internet telephony, Internet Phone to Internet Phone (PC2PC), Internet Phone to telephone (PC2Phone), telephone to Internet Phone (Phone2PC), Phone2Phone and Web2Phone. It needs the computer telephony gateway to connect computer and telephone as types of PC2Phone, Web2Phone and Phone2Phone. In this paper we proposed two computer telephony gateways that are easy to implement with the type of PC2Phone.

RTP is a transport protocol for real-time applications like voice and video transmission developed in 1996. The detailed description can found in Request for Comments 1889 (RFC 1889) issued by the organization of Internet Engineering Task Force (IETF). RTP is upon the User Datagram Protocol (UDP) that means the RTP packets sent to network without the mechanism of retransmission.

Then we introduce the quality measure of internet telephony and the spike detection proposed in 1994 and its solution of adaptive playout buffer. We experiment with the internet telephony software of SpeakFreely that can run in UNIX, Linux and Windows, and we got experiment data to prove the reason why packet loss and the spike. We offer a mathematical formula to measure the network bandwidth with RTP voice packets, then choose the optimal voice compression. We also offer an easy spike detection formula to early detect the spike and change to the more compact compression method at the beginning of the spike. Thus the voice quality of internet telephony is improved. Finally we have a discussion about the paper.

2.Two proposed structure of computer telephony gateway

Computer always added the telephony interface card to form a computer telephony gateway. The telephony card that is the key part of the internet telephony gathered the ability of modem, sound card. Relatively, the telephony card is so expensive that less people can afford it. And we will spare the modem, sound card when we used the telephony card. So we propose two structure of computer telephony gateway with the parts that personal computer already has.

First system structure we use any one of two modems to handle the call setup of PSTN and then the voice recording and playback by each modem. Second system structure we use a modem to handle the functions of PSTN and sound card to handle the voice processing.

2.1 Computer telephony gateway

The computer telephony gateway must handle the signaling and voice process of PSTN, and the PSTN system communicated with computer network by the computer telephony gateway. The system architecture of computer telephony gateway shows in Figure 2.1. It gives the signal of PSTN telecommunication system, such as dial tone, ringing, busy tone for call setup (call connection). And it detects DTMF tone from PSTN and decodes to check it is connected or not.

The computer telephony gateway also must be a D/A, A/D converter to translate analog signal to and from digital data. The voice compression method of G.729 to encode the voice data to a bit rate of 8 K bits per second was developed for efficient transmission for computer network in 1996. And the Real Time Protocol (RTP) was a transport protocol for real-time applications like voice and video transmission.

The issue of addressing, that translates the telephone number of PSTN and the IP address of the computer network, can be achieved by a matching table to match the telephone number and the IP address. And the PSTN used DC 48V and hybrid transmission to send and receive the voice with the same two telephone lines.

The computer telephony gateway always accomplished by a computer added a telephony interface card. The telephony card gave all the functions of PSTN and have D/A, A/D converter to convert the analog voice data to and from digital data and encoded and decoded the voice streams. So, the telephony card has the ability of modem to handle PSTN, of sound card to handle voice processing.

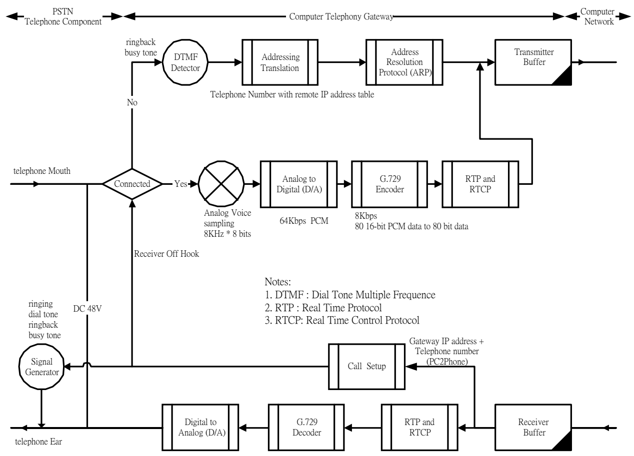

Fig. 2.1 The architecture of computer telephony gateway

2.2 Computer telephony gateway by two voice modems

A personal computer with a voice modem can be served as a telephone answering machine. The answering voice was pre-recorded, and answered when telephone call in. The functions of record and playback were processed not at the same time. We wanted to record and play back the voice simultaneously that was necessary for an internet telephony system.

With modem we can handle the PSTN signaling like call setup, ringing and voice processing of D/A, A/D conversion and voice compression. Two computers connected by internet or LAN and one computer remote login telephony gateway and dial a telephone number for call connection with remote PSTN telephone. After connecting, the sender computer sent voice from the microphone of its sound card and encoded voice for efficient transmission in internet. The encoded voice transmitted to telephony gateway, then gateway decode the voice data and playback the voice to PSTN by the software called "Voice Modem Control Program" (VMCP) in modem 1.

The telephony gateway recorded the PSTN voice also by the software called "Voice Modem Control Program" (VMCP) in modem 2. Then the gateway encoded the voice and transmitted the encoded stream to the receiver computer. The receiver computer played back the PSTN recorded voice by speaker. The system structure is as Figure 2.2 and the voice modems can be any internal modem card. If use the 8 ports multi-port modem, we can get a 4 line telephony gateway at most.

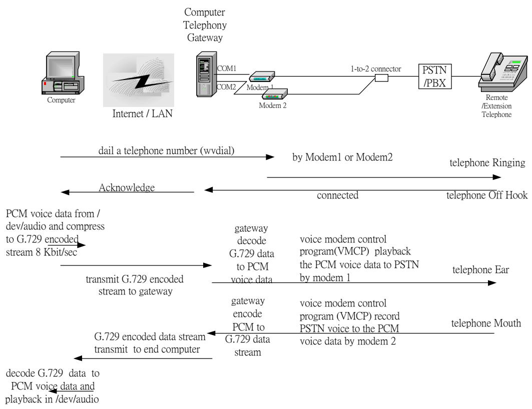

Figure 2.2 The system structure of computer telephony gateway by two voice modems

The voice modem at least is able to be a telephone answering machine. It's the first requirement of the voice modem. And the voice modem allows for entering transmit mode (VTX) to playback the voice and entering receive mode (VRX) to record the voice. We found almost all the modems disable the voice commands of playback (VTX) and record (VRX). Finally we found only the smartlink 5634PCV PCI modem card produced by Archtek telecom Corp. whose modem has enabled the IS-101 voice commands of the record and playback. There are two type of voice commands, the IS-101 voice commands and the Rockwell AT voice commands

2.3 Computer telephony gateway by a modem and sound card

The foregoing system used two modems that the software must use the technology of multithread to record and playback voice by each modem at the same time. It's not efficient for us. So we still use a modem to handle the functions of PSTN and use sound card to handle the voice processing.

Because the telephony card replaces all the functions of the sound card and the modem, as using the telephony interface card, the sound card and modem of a personal computer will be redundant. We will waste our sound card and modem when we use the telephony card. So we proposed another structure as Fig. 2.3.

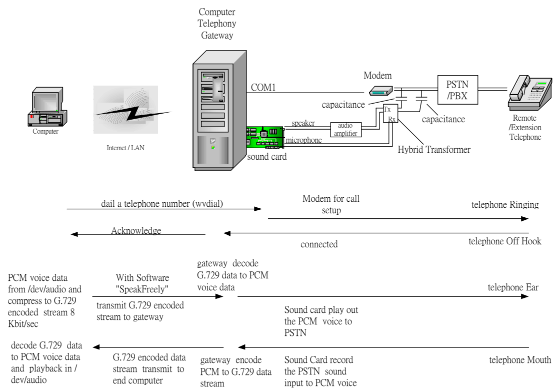

Figure 2.3 The system structure of telephony gateway by a modem and sound card

The sound card directly connects to the PSTN for voice processing in this telephony gateway. Two capacitances each serially connected to the two lines of PSTN for separating the DC 48V of PSTN from the sound card, but let the voice signal still transmit to and from PSTN. The speaker and microphone of the sound card totaled four lines must add a hybrid transformer that is a 4-to–2 connector for translating the telephone two lines that bi-directionally transmit voice signal to and from two lines' microphone and two lines' speaker.

Because the voice signal of the speaker line out is not loud enough, we can add an audio amplifier such as LM386 to amplify the voice signal, or we adjust the volume of the speaker by the software of the computer telephony gateway. We still remote login to the telephony gateway and use a modem to make a call connection, and then with the software of "SpeakFreely" the end computer can talk to remote telephone.

The necessary specific hardware and software for this structure are:

1. Electronic parts: We need two capacitances, an audio amplifier and one hybrid transformer to build a specific electronic circuit between PSTN and sound card.
2. Internet telephony Software: In this structure we can use any PC2PC software of internet telephony, like SpeakFreely, Netmeeting, iPhone, MediaRing, etc. Originally two computers can talk each other with duplex sound card by the PC2PC software. We used the functions of the PC2PC software and extended to a PC2Phone telephony gateway by adding a specific electronic circuit.

3. Quality measure and experiment results

In this section we experimented to found out the reason why the quality reduction in internet telephony with different voice compression and bandwidth.

3.1 Delay, jitter and spike

The delay, jitter and spike are the major index to measure the quality of internet telephony. The formula to computing the delay and jitter is as equation (1) (2), each RTP packets sent from sender every time period to the receiver side. S_i and S_{i+1} are the time that packet i and its following packet sent at the sender side, and R_i and R_{i+1} are the the time that packet i and its following packet received at the receiver side.

The delay of packet R_{i+1} and the jitter J_{i+1} of packet R_{i+1} are as equations (1)(2)

$$D = (R_{i+1} - R_i) - (S_{i+1} - S_i) \qquad (1)$$

$$J_{i+1} = J_i + (|D| - J_i)/16 = 15/16 * J_{i+1}/16 * |D| \qquad (2)$$

Where

D= the delay of packet R_{i+1}

J_i = the jitter of packet R_i

The spike issued 1994 in [3] is a situation that a very long time delay of packet and its following packets received in very short time interval. The spike is the real quality destroyer in internet telephony.

3.2 Experiment results

We transfer a voice file of 163840 bytes that is a 20.48 sec voice file between two computer with RTP protocol. First, we experiment with a PC by a 56 kbps modem that can connect to internet with PPP (Point to Point Protocol), which is a mechanism for creating and running IP (Internet Protocol) over a serial link using modems and telephone lines or using digital lines such as ISDN. That PC, located in Tao-Yuan runs the Linux operating system of

SuSE, connected to a SUN workstation runs Solaris in Hsin-Shu NCTU. We used the PPP protocol to connect PC and SUN workstation with a 56 kbps modem. Second, we experiment with the same two computers, but the PC was located in the campus of NCTU and used 10M ethernet card to connect the SUN workstation.

We transmit the voice file to the receiver computer and collect the arrival time and serial number and timestamp of each RTP packets in receiver computer. The SpeakFreely can choose any one of voice compression with PCM (no compression), ADPCM, GSM or LPC.

We added a short code as following in the isrtp function of "rtpaket.c" in SpeakFreely version 7.1 that function is in the receiver side to determine if a packet is RTP packet or not. If so, convert the RTP packet in place into a sound buffer. We then use the software of Excel to analyze the collected data and get the result of Table 3.1.

```
FILE *fw;        /*for file open and write */
struct timeb t;
ftime(&t);   /* get the system time */
fw= fopen("rtptimesize.txt","a");
/* to get the system time to the unit of millisecond,
and get the payload length , sequence number and
timestamp of each RTP packets */
fprintf(fw,"%ld.%#03d %d %u %u\n ", t.time,
                t.millitm, paylen, r_seq, r_ts);
fclose(fw);
```

It showed that the packet loss took place in insufficient bandwidth and the spike took place in network jam.

Table 3.1 The experiment result with 56 kbps modem and 10M campus network

	Voice Coding	Bit rate kbps	RTP Payload bytes	Sender packets time interval	Packet Loss /total	|Delay| Min/Max /Average (ms)	Jitter Min/Max /Average (ms)	Time used to send the voice file (sec)	Δtr : time interval of each received RTP packets	used /required band-width kbps
Low Band-width with 56kbps modem	No (PCM)	64	320	40ms	321 /511	10/130 /91	5.6/143 /120	24.536	40+40 ↑ ms	25.1 /78.8
	ADPCM	32	324	80ms	82 /255	20/2339 /77	6/208 /66	27.625	80+40 ↑ ms	24.3 /43.4
	GSM (no spike)	13.2	132	80ms	0 /256	0/38 /10	1/15 /10	20.358	80± 20ms	21.3 /20.6
	GSM (has two spike)	13.2	132	80ms	0 /256	0/1417 /30	1/109 /30	20.387	80± 40ms	28.9 /20.6
	LPC	5.6	56	80ms	0 /256	0/40 /0	2/8 /0	20.368	80± 40ms	13.2 /13.0
High Band-width With 10Mbps	No (PCM)	64	320	40ms	0 /512	0/39 /0	0/4 /0	20.435		79.5 /78.8
	GSM	13.2	132	80ms	0 /256	0/34 /1	0/5 /1	20.401		20.6 /20.6

The spike happen in GSM compression method is show as Table 3.2 that the normal value of Δtr_i is nearly 0.08 second, but the first spike happened in 107th packet and Δtr_i up to 1.497 second and next following 18 packets arrived very quickly. The next spike happened in 126th packet and the value of Δtr_i up to 0.595 second and next following 20 packets arrived very quickly.

Table 3.2 The spike happened in GSM compression method with 56 kbps modem

$packet_i$	arrival time	payload length	sequence number	timestamp	Δtr_i (sec.)
107	17.024	132	181151	34691321	**1.497**
108	17.034	132	181152	34691961	0.010
109	17.051	132	181153	34692601	0.017
110	17.074	132	181154	34693241	0.023
111	17.096	132	181155	34693881	0.022
112	17.118	132	181156	34694521	0.022
113	17.140	132	181157	34695161	0.022
114	17.185	132	181158	34695801	0.045
115	17.229	132	181159	34696441	0.044
116	17.251	132	181160	34697081	0.022
117	17.295	132	181161	34697721	0.044
118	17.318	132	181162	34698361	0.023
119	17.340	132	181163	34699001	0.022
120	17.362	132	181164	34699641	0.022
121	17.384	132	181165	34700281	0.022
122	17.406	132	181166	34700921	0.022
123	17.451	132	181167	34701561	0.045
124	17.473	132	181168	34702201	0.022
125	17.517	132	181169	34702841	0.044
126	18.112	132	181170	34703481	**0.595**
127	18.123	132	181171	34704121	0.011
128	18.138	132	181172	34704761	0.015

$packet_i$	arrival time	payload length	sequence number	timestamp	Δtr_i (sec.)
129	18.160	132	181173	34705401	0.022
130	18.183	132	181174	34706041	0.023
131	18.205	132	181175	34706681	0.022
132	18.249	132	181176	34707321	0.044
133	18.293	132	181177	34707961	0.044
134	18.338	132	181178	34708601	0.045
135	18.361	132	181179	34709241	0.023
136	18.404	132	181180	34709881	0.043
137	18.449	132	181181	34710521	0.045
138	18.493	132	181182	34711161	0.044
139	18.537	132	181183	34711801	0.044
140	18.560	132	181184	34712441	0.023
141	18.604	132	181185	34713081	0.044
142	18.648	132	181186	34713721	0.044
143	18.670	132	181187	34714361	0.022
144	18.715	132	181188	34715001	0.045
145	18.737	132	181189	34715641	0.022
146	18.766	132	181190	34716281	0.029

3.3 Network bandwidth measured with RTP packets

The voice data compressed or encoded to be the RTP payload and added RTP header, then added IP header, UDP header and ethernet header were transmitted in computer network as Figure 3.1

26 bytes	24 bytes	8 bytes	16 bytes	N bytes
Ethernet header	UDP header	IP header	RTP header	RTP payload

Figure 3.1 The length of each voice frames in RTP format

The bandwidth of the network needed to normally transmit voice data is as equation (3). And the time intervals of each sender packets are all the same as the time slot Δts.

$$BW = Br + (74*8)/\Delta ts \qquad (3)$$

where

BW = the bandwidth of the network
Br = the bit rate of voice packets transmitted
Δts = the time interval of sender packets
= (timestamp of sender $packet_i$-the timestamp of sender $packet_{i-1}$)/8000

We used the 56 kbps modem to transmit PCM and ADPCM voice that their bit rate is 64K and 32K separately. The voice streams of PCM and ADPCM needed the bandwidth of 78.8K and 43.4 K that 56 kbps modem can't afford. So bandwidth insufficiency resulted in the packet loss as Table 3.1. The experiment shows that a 56 kbps modem gives the speed about one half of 56 kbps bits per second for upload streams or download streams.

4. The adaptive playout buffer

In reference paper [3] the spike showed that almost 50 packets have not received during nearly one second. And many papers proposed adaptive playout buffer to solve the long delay of spike. The bit rate of the packets is 64 K bits per second and each packet size is 160 bytes that is 20 millisecond In reference paper [3]. We tested such algorithms of adaptive playout buffer in Table 4.1. It tells that the algorithms of adaptive playout buffer can't really solve the problem of the spike. The mechanism of adaptive playout buffer only tune the play out time and not suited for a long time delay. The average playout delay must be under 250ms that is considered as acceptable quality in internet telephony.

Table 4.1 The algorithms of adaptive playout buffer [3] to solve spike

Algorithms of adaptive playout buffer	Description	Packet discard (not in time to playout)	Gap silence (ms) ± 20ms	Playout delay=Playout time – transfer time	Remarks
Without buffering	The typical spike Packets playout time = packets arrival time	0	838	Max 959 ms Min 140 ms Average 343 ms	

Algorithm A	1.The first packet of a talkspurt: $p_i = t_i + \hat{}d_i + 4*\hat{}v_i$ The following packet j : $p_j = p_i + t_j - t_i$ 2.If (ni> di) then $\hat{}d_i = \alpha * \hat{}d_{i-1} + (1-\alpha) * n_i$ else $\hat{}d_i = \beta * \hat{}d_{i-1} + (1-\beta) * n_i$ $\hat{}v_i = \alpha * \hat{}v_{i-1} + (1-\alpha) * \lvert \hat{}d_i - n_i \rvert$, α=.998and β=.75	44 packets (packets 11-54)	990	Max 239 ms Min 239 ms Average 239ms	1.Too many packets not in time to playout, 2.reduce the average delay
Algorithm B	1.The first packet of a talkspurt: $p_i = t_i + \hat{}d_i + 4*\hat{}v_i$ The following packet j : $p_j = t_j + \hat{}d_j + 4*\hat{}v_j$ 2.If (ni < ^di) then $\hat{}d_i = \alpha * \hat{}d_{i-1} + (1-\alpha) * n_i$ else $\hat{}d_i = \beta * \hat{}d_{i-1} + (1-\beta) * n_i$ $\hat{}v_i = \alpha * \hat{}v_{i-1} + (1-\alpha) * \lvert \hat{}d_i - n_i \rvert$, α=.998 and β=.75	7 packets (packets 11-17)	770	Max 830 ms Min 208 ms Average 403ms	1.Reduce the Gap silence 2. have packet discard
Algorithm C	1. If spike then $\hat{}d_i = \hat{}d_{i-1} + n_i - n_{i-1}$ else $\hat{}d_i = \alpha * \hat{}d_{i-1} + (1-\alpha) * n_i$, α : 0.875 2. $\hat{}v_i = \alpha * \hat{}v_{i-1} + (1-\alpha) * \lvert \hat{}d_i - n_i \rvert$, α : 0.875 3. The first packet of a talkspurt: $p_i = t_i + \hat{}d_i + 4*\hat{}v_i$ The following packet j : $p_j = t_j + \hat{}d_j + 4*\hat{}v_j$	0	804	Max 1030 ms Min 165 ms Average 388 ms	1.Reduce the gap silence. 2.almost same with 4*^vj 3*^vj, 2*^vj
Proposed Method	1.Use the improved spike detection in section 5.3 and use the GSM compression method 2.In the 3rd time slot the packet is not received in receiver side, we change to the more compact compression method in sender side. (Because compression method developed in 1996 have much improvement in the spike issued in 1994)	0	<5*20+2*20=140		1.Decrease 80% gap silent

The adaptive playout buffer couldn't give any voice data when the long delay time nearly 1 second. The best way to solve the spike is to use more compact method of voice compression that reduces the probability of spike occurrence. Mentioned in section 5.3, to predict or detect in the beginning of the spike occurrence and immediately change the voice compression to more compact method as soon as possible.

5.Voice quality improvement in internet telephony

5.1 Detection of bandwidth sufficiency

From the experiment result of Table 3.1 we found that the arrival time interval of the packet i is one and a half bigger than the sender time interval as equation (4), and the arrival time interval of the following packet i+1 is also one and a half bigger than the sender time interval as equation (5), as the bandwidth of network is insufficient to transfer the voice packets.

$$\Delta tr_i > 1.5\Delta ts \qquad (4)$$

$$\Delta tr_{i+1} > 1.5\Delta ts \qquad (5)$$

where

Δtr_i = the time interval of the receiver $packet_I$ -$packet_{i-1}$

Δts = the time interval of sender packets

5.2 Detection of the spike

The phenomenon of bandwidth insufficiency is the successive longer time interval of each RTP packets, but the spike happens in a very long time interval of RTP packets and its following packets' time intervals is very short. An easy formula of spike detection is as following. The arrival time interval of the packet i is three times bigger than the sender time interval as equation (6). And the arrival time interval of the following packet i+1 is less than the sender time interval as equation (7).

$$\Delta tr_i > 3\Delta ts \qquad (6)$$

$$\Delta tr_{i+1} < \Delta ts \qquad (7)$$

where

Δtr_i = the time interval of the receiver $packet_I$ -$packet_{i-1}$

Δts = the time interval of sender packets

5.3 Solution of the spike

The better way to solve the spike is to use more compact method of voice compression that reduces the probability of spike occurrence. To predict or detect in the beginning of the spike occurrence in the preceding section and immediately change the voice compression to more compact method as soon as possible. The early you detect in the beginning of the spike occurrence, the less time loss for changing voice compression method.

The gap silence of spike will be three time slots (3Δts) plus the time of the sender received the message to change the compression method from the receiver side (Tm), then plus the network delay time to transfer the new compressed voice packets (Tn) and the processing time of new compression method in both sender and receiver side as equation (8). Tm is shorter than Δts because the message of Tm is much shorter than the sender packet's size of the original compression method. The network delay time to transfer the voice packets with the new and more compact compression (Tn) is less than or equal to Δts because as the voice packets are compressed compactly and it take less time to be transferred. The gap silence of the spike will reduce to 5Δts plus 2Tp as equation (8) to equation (10). The processing time of each compression method can refer Table 5.1 the compression methods and their features in section 5.4.

Gap silence of spike

$$= 3\Delta ts + Tm + Tn + 2Tp \qquad (8)$$

$$< 4\Delta ts + Tn + 2Tp \qquad (9)$$

$$\leq 5\Delta ts + 2Tp \qquad (10)$$

where

Δts: the time interval of the sender packets of the original compression method

Tm: the time of the sender received the changing message from the receiver side.

Tn: the network delay time to transfer the voice packets

Tp: the processing time of the new compression method

5.4 The algorithm for detection and solution of the spike or bandwidth sufficiency

By the section 3.3 we can measure the network bandwidth with RTP packets, or we know the network environment, we choose the optimal voice compression method by the Table 5.1. Then we detected the network's bandwidth and solve the spike and bandwidth sufficiency as Fig 5.1.

Table 5.1 The compression methods and their features

	Coding Type	Bit Rate kbps	RTP payload	Sender packets time interval	require band-width (kbps)	Process-ing time (ms)
G.711	PCM	64	320	40ms	78.8	0.125
G.726	ADPCM	32	324	80ms	39.4	0.125
G.728	LD-CELP	16	162	80ms	23.4	0.625
GSM	RPE-LTP	13.2	132	80ms	20.6	20
G.729	CSA-CELP	8	80	80ms	15.4	10
G.729A		8	80	80ms	15.4	10
G.723.1	ACELP	6.3	64	80ms	13.7	30
G.723.1	MP-MLQ	5.6	56	80ms	13.0	30

If the bandwidth of network is insufficient, we change to the more compact compression method to solve the packet loss and total transfer delay. And we use the detection formula to early detect the spike and change to the more compact compression method at the beginning of the spike. Thus the voice quality of internet telephony is improved.

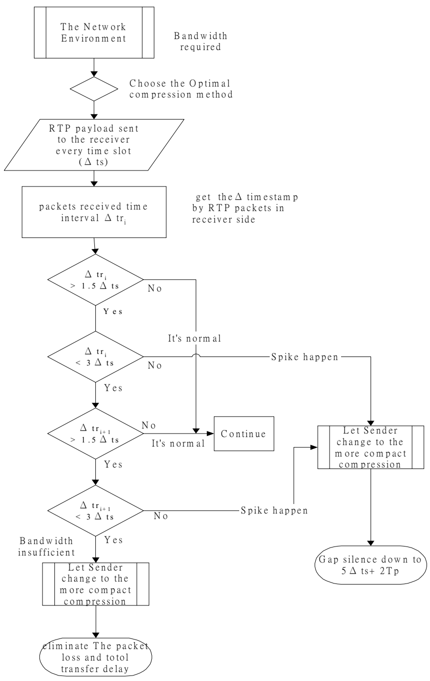

Fig.5.1 The algorithm for detection and solution of the spike or bandwidth insufficiency

6.Conclusion

We proposed two lower hardware cost of computer telephony gateways whose structures usually only had the ability of one telephone line of PSTN, not suit the trunk telephony gateway used for many telephone lines. But these two structures overcome the shortcomings of the conventional expensive telephony interface card.

The insufficient bandwidth of network resulted in packet loss. If we retransmitted the duplicated packets to reduce the packet loss, it will burden the loading of the network and get no benefit. The better way is to use more compression method to save the bandwidth and the packet loss will down to zero that the better voice quality was achieved.

In our experiment the probability of spike occurrence is about 25% with GSM compression by 56 kbps modem. In Table 3.2 it happened two successive spikes, the first gap silence is 1.497 sec and the second is 0.595 sec. With our issued algorithm, the gap silence will down to 0.460 sec. As we change the GSM to LPC compression that have 69% improvement for first spike and 22% improvement for the next spike.

7.Acknowledgements

Thanks for the friends worked in science-based industrial park gave me many technical advises for my innovations in computer telephony gateways.

Also I am appreciated of my advisor, Dr. Chen-Yi Lee, and my lab's classmates in the national Chiao-Tung uninversity, especially the Ph.D. candidate Ching-Che Chung, who gave me detailed explanations and enthusiastic assistance when I have any question.

Reference

[1] Ph.D. Sheng-Lin Chou, *Computer Telephony Integration System and Applications*, Lectures Notes for CSIE/NCTU Cource, Feb. 1999.

[2] "*RTP: A Transport Protocol for Real-Time Applications*", Request for Comments 1889, Network Working Group, IETF, Jan. 1996.

[3] R. Ramjee, J. Kurose, D. Towsley, and H. Schulzrinne, "Adaptive playout mechanisms for packetized audio applications in wide-area networks", INFOCOM '94, Networking for Global Communications, 13th Proceedings IEEE, vol.2, pp.680 -688 , 1994.

[4] J.C. Bolot, S. Fosse Parisis, and D. Towsley, " *Adaptive FEC-based error control for Internet telephony*", INFOCOM '99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies Proceedings, IEEE Vol. 3, pp.1453 - 1460, 1999.

[5] Craig Hunt, TCP/IP Network Administration, Second Edition, O'REILLY & Associates, Inc., Dec. 1997.

[6] S. Rudkin, A. Grace, and M.W. Whybray, "Real-Time Applications on the Internet", BT Journal Vol. 15, No. 2, April 1997.

[7] "Coding of Speechat 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction(CS-ACELP)", ITU-T Recommendation G.729, March 1993.

[8] Corwin Light-Williams, Joshua Drake, "*the Linux PPP-HOWTO*", July 2000, http://www.linuxdoc.org/HOWTO/PPP-HOWTO/index.html