How to strengthen the identity identification and communication data non-repudiation between both communication sides of VoIP (Voice over Internet Protocol) has undeniably become two very important topics in the development of VoIP. In this study, without the modification of the current communication architecture, RTP hidden channel that has error resilient mechanism is implemented to carry identity-carrying information so as to solve identity identification and non-repudiation issues. Error resilient mechanism can surmount implicit data loss due to RTP packet loss. As compared to related research, this study, with its lower cost, can provide VoIP user a selectable identity identification safety mechanism.

1. Introduction

VoIP based on TCP/UDP, due to its inheritance of the disadvantages of basic communication protocol, will have to suffer the threats of identity fraud, Man-in-the-Middle attack, etc.[7]. The CIO (Chief Information Officer) of Royal Mail, England has warned that inappropriate use of VoIP application software can possible let the enterprise expose under the hacker attack and the threat from malicious program, it might even possibly the important confidential information of the company [4]. According to 10 forecasts made by IDC (2005) on the information telecommunication for Asia Pacific region (excluding Japan) for year 2006, it is found that identify identification and access are going to be the major focuses [6].

The development and growth of entire VoIP will be brought up by the enterprise sector, hence, how to ensure communication non-repudiation through identity identification and communication data binding will be another challenge of the commercial development of VoIP [12].

1.1 Research background

Among VoIP related communication protocols, SIP (Session Initiation Protocol) as prepared by IETF occupies a largest percentage. SIP is in charge of the starting of call signaling, and the media message streaming propagation is done by RTP (Real-time Transport Protocol). Until now, studies regarding identity identification in VoIP all focus on the use of SIP communication for handling those kinds of problems. However, in this study, RTP communication protocol will be used to solve simultaneously identity identification and non-repudiation problems. In the study, it is assumed that the current Digital Certification mechanism is perfect, and the legal issue of the user identity identification won’t be investigated in depth here.

1.2 Research motivation and objective

Currently, VoIP is mainly based on SIP/RTP as its basic structure, but since no encryption has been performed on SIP packet, identity and certification information in the communication process can be easily detected by network packet searching tool [7]. SIP communication process might pass many different mainframes and equipments; hence, there are still many standardization topics to be surmounted in order to achieve End to End safety mechanism in the internet. Although there are many researches showing the progress of identity identification through RTP, yet the use of UDP for RTP packet transmission might lead to packet loss [3], which in turn might the problem of incomplete transmission of hidden information.

This study is based on a background of SIP/RTP VoIP; without the change of communication protocol, RTP with error resilient information hiding technique is realized to carry information that is useful for identity identification so as to solve identity identification and non-repudiation issues. The use of Reed Solomon code error resilient mechanism can surmount the problem of hidden information loss due to packet loss.

2. Literature review

The technical fields investigated in this study include VoIP basic communication protocol, SIP safety mechanism, QoS, Information Hiding and Error Correction technique.

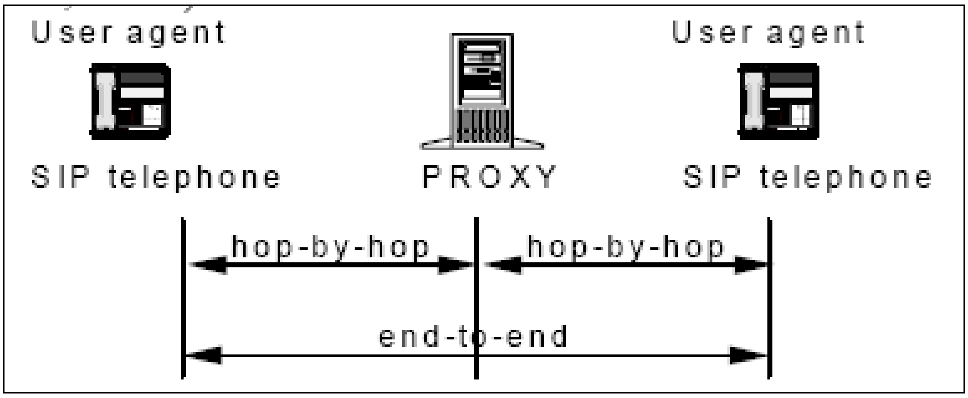
2.1VoIP basic communication protocol

Among VoIP related communication protocols SIP communication protocol has gradually replaced the initially developed H.323. 3GPP has pointed out that SIP will be the next key internet communication protocol [8]. After SIP has set up the initial connection between both sending and receiving sides, RTP will be used to achieve voice image transmission. RTP is used for the transmission of media information, and timestamp is used to achieve synchronization ability; finally, serial number is used to decide if there is any packet loss, etc.

RTP can be used to transmit information of normal format or defined format. PCM standard G.711 as made by CCITT in 1972 is a widely used VoIP format with a speed of 64kbit/s, which is equivalent to the non-linear quantified μ law or A law of 12 bits linear quantification.

2.2 The development and analysis of VoIP safety mechanism

It was pointed out in a study performed by Rosenberg (2002) that clear text SIP message will suffer from the issues of Spoofing, Hijacking and message correction. It was also pointed out in a research of Geneiatakis (2005) that malicious SIP message will lead to illegal access [8]. SIP safety structure can be divided into two types, namely, End to End and hop by hop [10]. The use of SIP to set up communication stage protection is currently the core development issue and direction of VoIP.



SIP is a communication protocol under development, which is especially true for security topic. After the settlement of SIP RFC3261 in 2002, many security application standards are gradually produced. In the proposed safety mechanism, relative safe selection projects have been provided, but some real topics are hidden inside. Arkko(2003) had proposed that the user should have the capability to select the needed safety mechanism according to his/her own need [8]. Analysis on the related solution projects as proposed by SIP is as in the following:

(1) HTTP Digest authentication: Ferguson and Schneier (2003) had pointed out that such mechanism might suffer from man-in-the-middle attacks, DoS.

(2) IP sec: In internet application, serious overhead will be generated.

(3) TLS and SIPS: The lack of PKI facility might lead to the realization barrier on internet.

(4) S/MIME: The same as IPsec, overhead will be the result; meanwhile, because of the lack of PKI facility, realization barrier on the internet will be formed.

From the above analysis, it is found that the use of these mechanisms has the hidden cost consideration of the realization of End to End safety channel; it is still in the development stage for many international giants in VoIP and safety solution projects, for example: international giant equipment manufacturer Juniper has proposed individual project by targeting at the individual need of an enterprise, but for Nortel, it puts its focus on Intranet and VPN.

2.3 QoS

For the QoS evaluation of IP voice, it can be discussed in two aspects; objectively, it includes four directions, which are respectively, time setup, packet loss rate, Jitter and point to point delay (latency)[3]. Subjectively, MOS [3] can be used for the evaluation.

In this study, factors affecting QoS include:

(1) When UDP is used for the transmission of RTP packet, packet loss problem might affect the voice transmission quality.

(2) Information hiding will destroy the original voice data; in the mean time, the operation time for the realization of information hiding might further lead to RTP packet delay.

2.4 Information hiding

Through information hiding technique, Authentication Information can be hidden to the target object for protection. Two important characteristics of information hiding are: non-perceptual and robustness.

(2006), least significant bit (LSB) method is used to hide the information to be hidden to the least bit of the sampling point of G.711 voice data. When MOS (Mean Subjective Scale) is observed to evaluate the effect of this method on voice quality, it can be found that human ear can not recognize the difference after change [11]. This method can suppress the exaggeration of the original voice error to achieve the purpose of information hiding.

2.5 Error detection and correction

For the resilience of error, it is usually needed to add redundant data so that the destroyed data or lost data can be recovered. Reed Solomon code has the compensation function for packet or bit loss; hence, in the early stage, it is seen as standard for error correction code and is applied in many storage systems for the processing of burst [13]; another famous application is PDF417 two dimensional barcode [4]. PDF417 can read back complete data from damaged barcode (Moore, 1994); its flexible fault tolerance can provide different resilient power to as high as 50%.

Reed Solomon principle divides the data into a block of fixed size and the coding is done through the coefficients a polynomial represented by finite field [4]. The block of Reed Solomon code is represented by (n, k), n represents symbol number of each block after coding, k is the original data of each coded block, t means the maximum number of errors that can be handled, and the formula is n = k + 2t. The decoding part can be represented as r(x) = v(x) + e(x), r(x) means the message received at the reception end, v(x) means the original coding of Reed Solomon code, e(x) is the noise; the error correction code contained in v(x), within tolerable error number, can help to get back to correct information. r(x) error can be divided into two types: E Error (Erasure Correction), which is caused by v(x) information loss and it is a type of known position error. T Error (Error Correction): e(x) will lead to v(x) to contain error information; this type of error resilience will need to find out data error position and to find out correct solution.

3. Research design

In this study, Reed Solomon error correction function is used to realize RTP module with information hiding technique. This model is used to verify the resilient effect after information has suffered packet loss. Moreover, the feasibility of this study structure is further evaluated and proved.

3.1 Experimental system structure

Figure 2 is the system structure of this study. The request end will convert the sampling voice signal first into G.711 code. At the same time, in order to hide the information, division is needed according to the data block size of Reed Solomon code, then correction code coding and conversion into Bit Stream on each block need to be performed; finally, LSB method is used to hide them sequentially into RTP packet (Figure 3). After the request end has received RTP packet with hidden information, in one aspect, it will output the packet into voice message, in another aspect, it will withdraw the hidden information and fill them back sequentially into each Reed Solomon code block; facing with the situation of finite packet loss and noise interference, error correction code is used to perform the resilience decoding of error information so as to get back the lost or error information.

In system realization, WinRTP(V2.1) module and Schifra Reed-Solomon ECC Library(V0.01) are going to be used to complete the setup of system model.

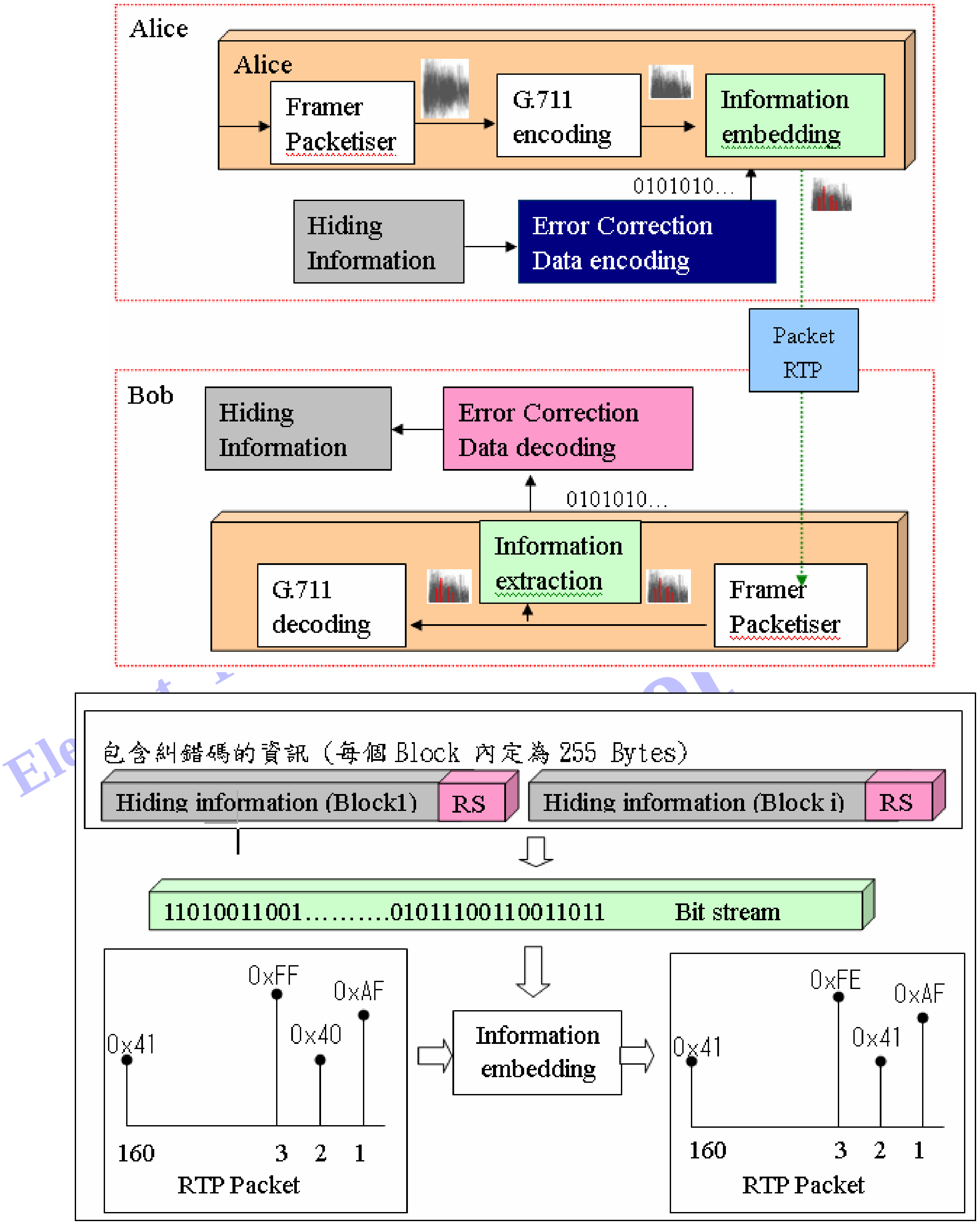


Figure 1 Illustration of information hiding method

3.2 System communication mechanism

In the research system, a mechanism for identity identification information exchange and communication between both communication sides should be realized so as to complete the identity identification operation. Between both communication sides, mutual transfer of control information, parameter and identity identification information as in table 1 is going to be done to complete identity identification operation handshake; the message communication process is as in figure 4, both end points perform control signal exchange through the designed protocol so as to activate identity identification.

Table 1 Communication mechanism for identity identification operation

Request end	Requested end	Transmission code
Certification request		Control information: CREQ
	Transmission operation parameters	Control information: CSTR Experiment related parameters.
Reception parameter		Control information: CACK
	Transmission of data	Control information: D Hidden information length (n)
Information reception		Control information: CDAT
	Transmission completed	Control information: CEND
Certification failure		Control information: CROK
Certification success		Control information: CRNG

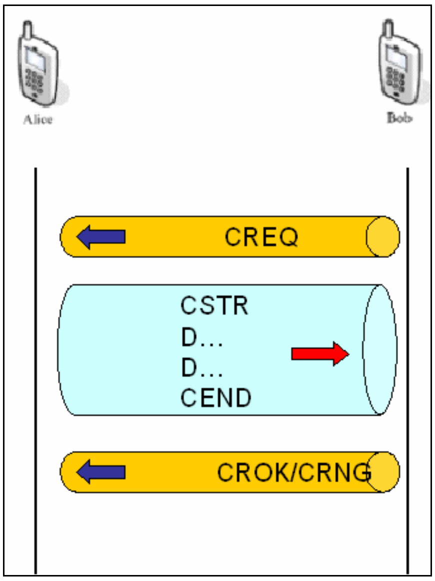


Figure 2 Delivery handshake procedure for hidden information

3.3 Experimental parameters

Table 2 describes the related parameters for the experiments performed on this research system. The experiment is going to use RTP Payload G.711 voice information as background to perform identity identification information hiding. Each RTP packet contains 160 sampling points and 50 packets are going to be transmitted each second. Since for G.711 format, when the packet loss rate reaches 10%, the voice quality is at the acceptable lowest limit, hence, this experiment is going to use 10% packet loss rate as the limit. In the experiment, LSB method is going to be used to realize information hiding to transmit 1024 Bytes of information in fixed way so as to adjust the error correction code length and the hidden information length of each packet in order to evaluate the performance of the experimental system.

Table 2Experimental parameter information

Experimental parameter	Selectivity
RTP Codec	G.711 A law
RTP length	20ms
Reed Solomon code block length	255 (Bytes)
Information hiding method	1 (Bits):LSB
Amount of data transmitted	1024 Bytes
Limit value of packet loss rate	10%
Error correction code length	16,32,64(Bytes)
Information hiding length	5,10,15(Bytes)

3.4 Evaluation keys of the research system

In the system experiment, the evaluation elements that need to be evaluated include:

(1) Voice quality: Information hiding algorithm might generate delay on voice processing. The hidden information damage in the voice information content might affect the user’s perception on the voice quality. Evaluation on the subjective and objective voice quality change caused by voice information hiding will be helpful in understanding the feasibility of the experimental method.

(2) Cost element: In the study, information hiding method is used for the transmission of identification information; due to the capacity limitation on the hidden information, time cost needs to be used to gain the safety. In the experiment, the transmission time variances caused by different transmission environments will be understood and the result will be provided to the user for a comparison of the cost needs to be consumed in the selection of different safety mechanisms. According to the experimental parameters, it can be predicted that the transmission time for the fixed transmission of data amount of 1024 Bytes in this experiment is as in table 3.

(3) Resilient performance: Different error correction grade will not only affect data transmission length but also the degree of packet loss that can be taken by this experiment. (Table 4)

Table 3 Information hiding transmission time forecast

Error correction code length (Bytes)	Amount of single packet information hiding (Bytes)	Number of block transmitted	Packet number in a single block	Forecast time (ms)
16,32	5	5	51	5100
16,32	10	5	26	2600
16,32	15	5	17	1700
64	5	6	51	6120
64	10	6	26	3120
64	15	6	17	2040

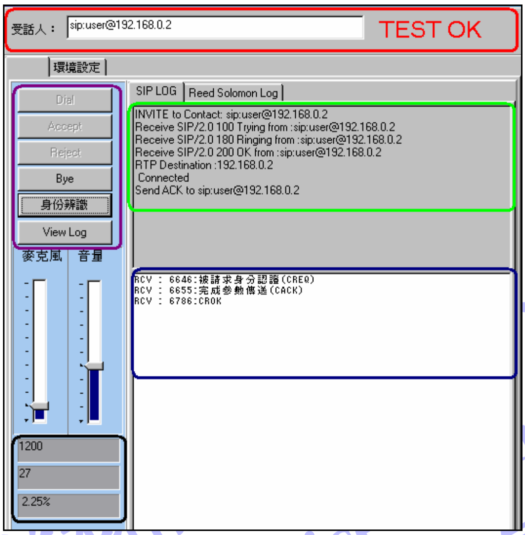
Table 4 Packet loss threshold forecast that can be taken by different error correction level

Error correction code length Bytes	Information hiding length Bytes	Amount of data transmitted Bytes	Single block packet number (pc)	Threshold value of lost packet (pc)	Packet loss rate that can be taken
16	5	1275	51	3	5.9%
16	10	1275	26	1	3.8%
16	15	1275	17	1	5.9%
32	5	1275	51	6	11.8%
32	10	1275	26	3	11.5%
32	15	1275	17	2	11.8%
64	5	1530	51	12	23.5%
64	10	1530	26	6	23.1%
64	15	1530	17	4	23.5%

4 Experiment and evaluation

4.1 Demonstration of system realization

In this study, system realization is used to test the performance platform. Feasibility and optimized threshold value test is done through this performance platform; please see figure 5 for the main experiment operation area of the system, as can be seen, the red block means the SIP URL of the receiving end and identity identification operation result. The green block means SIP communication setup process, which includes INVITE, 100 Trying, 180Ringing, 200 OK and ACK. The blue block means the received process control information which includes information such as CREQ, CACK and CROK. The operation zone is as in the purple block, which can be used for Dial, Accept, Reject, end a dialogue (Bye), identity identification and View Log. The black block on the lower left side is used to display the statistical data of the received packet; as can be seen in figure 5, this system has received 1200 RTP packets, lost 27 packets, with a loss rate of 2.25%.



受話人：sip:user@192.168.0.2

TEST OK

```
INVITE to Contact: sip:user@192.168.0.2
Receive SIP/2.0 100 Trying from :sip:user@192.168.0.2
Receive SIP/2.0 180 Ringing from :sip:user@192.168.0.2
Receive SIP/2.0 200 OK from :sip:user@192.168.0.2
RTP Destination :192.168.0.2
 Connected
Send ACK to sip:user@192.168.0.2
```

```
RCV : 6646:被請求身分認證(CREQ)
RCV : 6655:完成參數傳送(CACK)
RCV : 6786:CROK
```

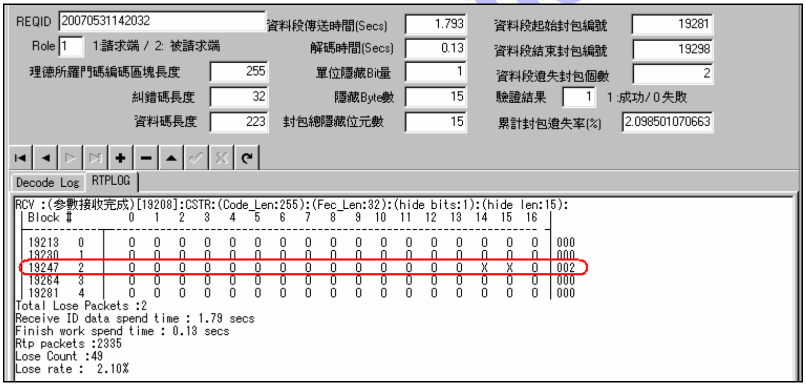
1200

27

2.25%

Figure 3 A screen showing the experiment system operation area

The statistical function of experimental log can be used to record the experiment related settings and parameters and performance data; as shown in figure 6, the number 14 and 15 packet of the third Reed Solomon code block (Block # 2) in this experiment is lost, but this is does not affect the subsequent identification because the error correction code in Reed Solomon block can provide resilient power to recover 30 Bytes of data in two lost packets. The data transmission time is about 1.793 seconds.



REQID 20070531142032

Role 1 1:請求端 / 2: 被請求端

理德所羅門碼編碼區塊長度 255

糾錯碼長度 32

資料碼長度 223

資料段傳送時間(Secs) 1.793

解碼時間(Secs) 0.13

單位隱藏Bit量 1

隱藏Byte數 15

封包總隱藏位元數 15

資料段起始封包編號 19281

資料段結束封包編號 19298

資料段遺失封包個數 2

驗證結果 1 1 :成功/ 0 失敗

累計封包遺失率(%) 2.098501070663

```
RCV :(參數接收完成)[19208]:CSTR:(Code_Len:255):(Fec_Len:32):(hide bits:1):(hide len:15):
 | Block #          0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16 |
 -------------------------------------------------------------------------------------
 | 19213   0  |     0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0 | 000
 | 19230   1  |     0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0 | 000
 | 19247   2  |     0   0   0   0   0   0   0   0   0   0   0   0   0   0   X   X   0 | 002
 | 19264   3  |     0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0 | 000
 | 19281   4  |     0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0 | 000
Total Lose Packets :2
Receive ID data spend time : 1.79 secs
Finish work spend time : 0.13 secs
Rtp packets :2335
Lose Count :49
Lose rate :  2.10%
```

Figure 4 A screen showing the statistical function of experimental log

4.2 Platform test and performance evaluation

The goal of this study is to set up an information hiding channel to the user as a selectable safety mechanism. In order to achieve this goal, related factor evaluation that needs to be done and experimental results are as in the followings:

4.2.1 Voice quality evaluation

In the objective voice quality evaluation, Fluke Network Protocol Expert voice quality measurement tool is used to perform voice quality experimental data collection on comparison group (no identity identification implemented) and experiment group (identity identification implemented) of two sets of end points in the local area network and wide area network so as to understand the effect of research system on voice quality. It can be proved from table 5 that this research system does not have significant effect on the voice quality (delay), and the voice quality evaluation index (MQS, R Factor) is also within acceptable range (MQS > 3.6, R Factor > 70).

Subjective voice quality evaluation is used to understand, after the implementation of information hiding technique, the subjective perception of the user on voice quality. In this experiments, 10 users are doing MOS evaluation on the voice after the adoption of LSB method; it can be seen from the evaluated data of 3.94 (MOS 良:3.5~4) that the LSB method adopted in the experiment only has limited effect on the subjective quality of the voice.

Table 5 Objective voice quality evaluation data table

	LAN		WAN	
	Comparison group	Experimental group	Comparison group	Experimental group
R Factor	93	93	88	87
MQS	4.18	4.17	4.04	4.03
Packets(Rec.)	6208	5483	6979	6673
Lost Rate	0.00	0.00	1.8%	1.7%
Delay (ms)	3.74	4.97	65.02	43.39
Jitter(ms)	2.28	6.16	4.16	6.17

4.2.2 Evaluation of system transmission cost

Two environments of local area network and wide area network (campus network versus 8M/512K ADSL) are used for the system transmission cost evaluation on different experimental combinations; through a comparison between time measurement table of table 6 and forecast value of table 3, it can be seen that on two end points of local area network, the transmission cost difference is about 6%. For wide area network, the two end points and the forecast value has a difference amplified to about 25%~46%.

From the voice quality experiment, it can be seen that wide area network has more delay as compared to local area network; the delay factor leads to that wide area network will need more time to realize identity identification and communication mechanism; therefore, wide area network has higher transmission cost as compared to local area network.

Table 6 Measurement table for identity identification time (unit: ms)

error correction code length Bytes	Information hiding length (Bytes)					
	5 Bytes		10 Bytes		15 Bytes	
	LAN	WAN	LAN	WAN	LAN	WAN
16	5402	6540	2774	3601	1811	2432
32	5402	6806	2748	3528	1807	2483

64	6441	7653	3276	4102	2157	2796

4.2.3 The resilient performance of transmitted information

In the resilient performance experiment, dual way identity identification experiment will be done through campus network (100M) versus ADSL (8M/512K) broadband network. Different combinations will be based on information hiding length (5,10,15 Bytes) and error correction code length (16,32,64Bytes), and each set of experiment will be with performed 50 times of identity identification operation. It is found in the data collection through Fluke Network PE voice quality tool that under 512K [the certified end]->100M environment and 16 Bytes error correction code, the 20 minutes packet loss rate is 2.3%, with a sudden highest loss rate of 3.91% (30 秒); within 450 times of identity identification operation, 260 (58%) times of packet loss was encountered, through error correction code mechanism, 144 times of packet loss can be recovered; through the use of 32 Bytes error correction code, the information resilient success rate is at least of 78% (The hidden length is 15 Bytes); through the use of 64 Bytes error correction code, the information resilient success rate is at least 90% (The hidden length is 10 Bytes). Under 100M[the certified end]->8M environment, 20 minutes packet loss rate is 1.7%, and the sudden highest loss rate is 3.91% (30 seconds); in 450 times of identity identification operations, 204 (45%) times of packet loss were encountered, then through error correction code mechanism, 200 times of packet losses can be recovered; through the use of 16 Bytes error correction code, the information resilient success rate is at least 96% (The hidden length is 10,15 Bytes); through the use of 32, 64 Bytes error correction code, information resilient success rate of 100% can be realized. For the related data, please refer to table 7.

Table 7 A statistical table for the resilient success rate of identity identification information (50 times each)

Experimental environment	Error correction code length	Information hiding length (Bytes)		
		5	10	15
512k -> 100M	16	38.0%	32.0%	64.0%
	32	92.0%	84.0%	78.0%
	64	96.0%	90.0%	94.0%
100M -> 8M	16	100.0%	96.0%	96.0%
	32	100.0%	100.0%	100.0%
	64	100.0%	100.0%	100.0%

5. Conclusion and future research and development

It can be seen from the experimental evaluation result that after the adding of error correction code and information hiding mechanisms in the RTP module of the research system, the realized hidden channel will then have error resilient function, and the packet loss issue in the transmission process can indeed be surmounted; under the environment of lower broadband, the use of 64 Bytes error correction code can let the success resilient rate reach more than 90%; under higher broadband environment, the use of 32 Bytes error correction code can let the success resilient rate reach 100%. The experiment also proves that this research method does not have significant effect on the voice quality. Therefore, this research method can provide a selectable safety mechanism so that through hidden information transmission, the user identification and non-repudiation value can be reinforced.

This study also triggers some topics which can be used as reference for future research and development:

(1) Algorithm improvement: It can be seen from an analysis on the resilient failure data (figure 7) that the sequential hidden algorithm adopted in this research has insufficient loss bearing force for continuous packet, and longer error correction code needs to be used to enhance the packet loss threshold of single block (yellow line); therefore, how to improve the algorithm, to diverge the packet loss risk, and to let the research structure not be limited by block loss packet threshold so that fewer error correction code needs to be transmitted to achieve better resilient rate are thus very important; assume the block number of Reed Solomon code is i, n is the block number that needs to be transmitted, Zi(b,f,h) is the loss packet number in single block, and each block length is represented as b, error correction code length condition is represented as f, hidden information length is represented as h, the loss packet threshold number of each block is represented as P(b,f,h), then the resilient failure formula generated by the current algorithm can then be represented as (equation)1; as long as there is one block packet loss number larger block loss packet threshold value, it will then cause resilient failure. Equation can be reviewed and improved as in (equation) 2, if this conditional equation can be met, then under narrower bandwidth of the system, the expected identification success rate can be further enhanced; the related algorithm improvement is a topic that can be investigated further in depth by this research.

$$MAX_{i=0}^{n-1}(Z_{i(b,f,h)}) > P_{(b,f,h)} \qquad \text{(式) 1}$$

$$\sum_{i=0}^{n}(Z_{i(b,f,h)} > \sum_{i=0}^{n}(P_{(b,f,h)})\ ) \qquad \text{(式) 2}$$

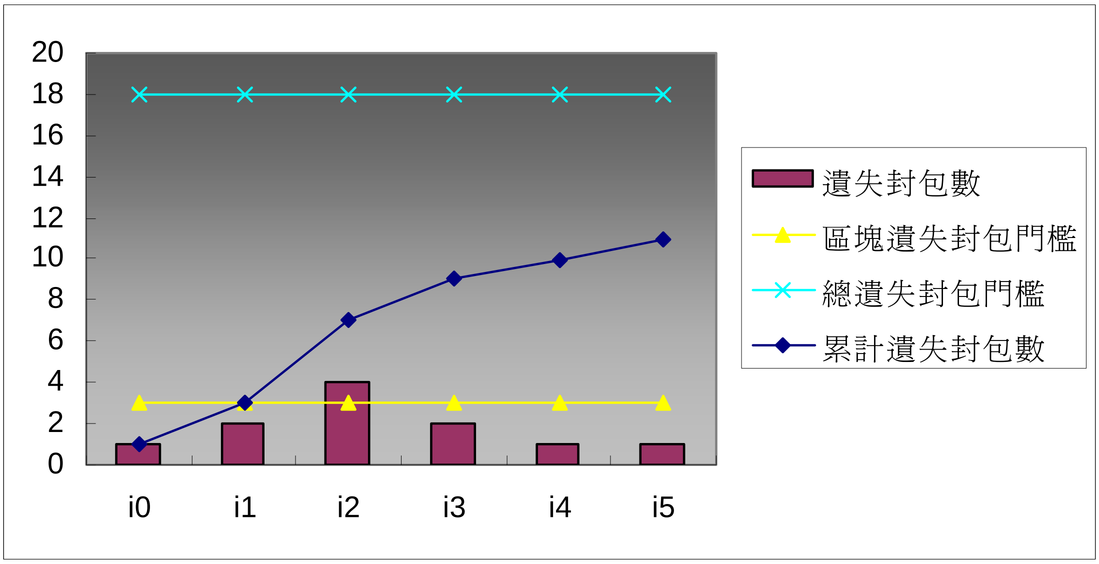


Figure 5 Illustration of the information hiding transmission failure threshold, Reed Solomon code block (255,191)

(2) Complete structure study: This study is mainly focused on the investigation of hidden channel operation,

there is no subsequent identity identification mechanism realized in the structure; in the future, it is hoped that this topic can be added so as to supplement the research completeness.

(3) Multi-element network environment test: In the resilient performance experiment of this study, single wide area network is used to complete data collection; but facing with the current multi-element network environment, it is hoped that in the future, more different network environments can be used to measure factors that might affect the performance so that the structure of this study can be applicable to different network environments.

Acknowledgement

This work was supported in part by TWISC@NCKU, National Science Council under the Grants NSC 96-2219-E-006-009.