ATLANTIS PRESS

Advances in Computer Science Research, volume 44
3rd International Conference on Wireless Communication and Sensor Network (WCSN 2016)

Dynamic Key Updating With Voice Quality Preserving for Voice over Internet Protocol Based on Steganography

Li-Ping Zhang
College of Computer Science, China University of Geosciences, Wuhan, 430074, China
carolyn321@163.com

Shan-Yu Tang
College of Computer Science, China University of Geosciences, Wuhan, 430074, China

Yi-Jing Jiang
College of Computer Science, China University of Geosciences, Wuhan, 430074, China

***Abstract*—Voice over Internet Protocol (VoIP) achieves cheap cost and provides flexible features by delivering voice packet over Internet Protocol. Compared with conventional public-switched telephone network, transmitting voice packets over public channel has to face more challenges due to the security threats and quality loss. To protect the voice packets transmitted over unsecured Internet, a shared key should be updated during a VoIP call. On the other hand, VoIP is more sensitive to transmission latency, so the flooding message traffic from key updating may cause the degradation of speech quality. This paper presents a dynamic key updating scheme without affecting the quality of VoIP calls. Unlike previous works, steganography technology is employed with encryption to provide protection of transmitted information in an unsecure channel. Moreover, in the proposed scheme, the quality of the VoIP call is preserved by adopting steganography technology to avoid the flooding message traffic. The security analysis demonstrates that the proposed scheme achieves the secure dynamic key updating without sacrificing the quality of VoIP calls.**

Keywords-VoIP; key updating; Steganograpy; AES

I. INTRODUCTION

With the rapid evolution of packets witched network technology, there is a growing trend in real-time voice communication using Internet Protocol (IP). Voice over Internet Protocol (VoIP) is a technology that delivers voice packets over an IP network. Different from traditional Public Switched Telephone Networks (PSTNs), in VoIP, voice is compressed and converted to digital voice packets, then the voice packets can be transported over IP network to achieve low cost and flexibility implementation. Since the cost for establishing and operating telephony services is reduced significantly, the prospects of VoIP applications are becoming promise to replace the traditional PSTNs. Furthermore, VoIP is easy to integrate with audio and video on VoIP platforms, so it can provide more flexible features than PSTN services. Various VoIP communication products such as Skype, Google Talk are already applying on the Internet, which provide cheap and even free phone calls with good quality. There are three core protocols for VoIP, Session Initiation Protocol (SIP), Session Description Protocol (SDP) and Real-time Transport Protocol (RTP). SIP is a signaling protocol used to establish or terminate a session between communication participants [1]. SDP as a company of SIP is applied to negotiate the types of media sessions [2] and RTP is aimed to deliver voice packets based on an established session [3].

However, different from transferring packets for texts and pictures, transmitting voice packets over Internet has to face more challenges. Although performance issues such as packet loss and jitter may be solved to afford the VoIP applications over Internet by using the broad-bandwidth technologies, VoIP communication is still experienced by the problem of quality loss. Furthermore, VoIP infrastructures are deployed in an open network environment, so the voice packets running over the Internet could be eavesdropped simply. And the privacy and value information could be compromised by several attacks [4] since voice packets transmitted over VoIP are unencrypted.

To provide the security of voice communication, a secret key should be shared between the caller and called to encrypt/ decrypt the voice packets over Internet. However, using one shared session key for a long time is not secure enough to protect the VoIP voice data throughout the whole VoIP call session [5]. Brute force attacks may be performed successfully due to the increase in computational power. Therefore, the shared session key should be updated frequently during a VoIP call to resist brute force attacks. The secure negotiation of a shared key has been widely researched and some key agreement schemes achieve security with good performance [6]. However, few works have been done to solve the problem of dynamic key updating for VoIP networks. Two approaches are widely used to realize the key updating, one is restarting the key agreement scheme to generate a new shared key, and the other is updating the shared key by using previous shared keys without extra communications. However the method of restarting key agreement process frequently would cause the flooding message traffic seriously and make users cannot tolerate the degradation of speech quality. So, this solution is inappropriate for dynamic key updating in VoIP

environments. In the local key updating method, no interactions are needed in the rekey process, so this approach does not affect the quality of the VoIP call. But, the relationships between the new shared key and the previous shared keys make this solution subject to some attacks. Once the adversary compromises a shared key, she/he can compromise other shared keys by using the links between the shared keys. Moreover, the synchronization of the key updating is another tractable issue in this solution. Therefore, the local key updating method is also unsuitable for VoIP and new approaches to realize the dynamic key updating for VoIP are sought.

In general, the previous works of key updating for VoIP systems are based on cryptography relying on some intractable mathematical problems [7]. Although various cryptography based schemes are proposed [8, 9, 10], key updating in VoIP is still an unsolved issue. In order to achieve frequent key updating for VoIP from both security and performance viewpoint, we propose a novel key updating scheme based on steganography and cryptography. In seganography, the secret message transmitted in an unsecure communication channel is not "unreadable" but "unobservable", which means that any other except the intended recipient can neither observer the existence nor know the content of the secret information [11]. Steganography hides the existence of the information to obfuscate the fact of communication rather than alters the transmitted data. Compared with cryptography, the goal of steganography is to conceal the existence of the secret message while the aim of cryptography is to conceal the content of the secret information. Advanced Encryption Standard (AES) as a specification of standard symmetric encryption has been widely adopted. Compared with other encryption algorithms, AES preserves the merits of both speed and security. It has been applied to provide secure end-to-end VoIP service. For example, a 256-bit AES encryption mechanism is adopted to encrypt/decrypt voice data in Skype. Different from previous works, in our scheme, Advanced Encryption Standard (AES) is adopted to encrypt the key materials and steganography technology is employed to hide the encrypted key materials.

The rest of this paper is organized as follows. Section 2 describes the secure and dynamic key updating scheme in detail. The security analysis is presented in Section 3. And the experiments and performance results are shown in section 4. The paper is concluded in Section 5.

II. SECURE AND DYNAMIC KEY UPDATING OVER VOIP NETWORKS

Cryptology algorithms are widely adopted to protect the real-time voice communications which rely on encryption keys. The cryptology algorithms are public and known by anyone including the adversary whereas the keys are kept secretly only known by the communication parties. Therefore, the security of the communication depends on the encryption keys. However, the encryption key may compromised from the brute force attack due to the usage of a single encryption key in a long time over VoIP communications. To resist such attacks, key updating mechanism should be provided to replace the current key with a new one after a time period of VoIP calls on the public Internet to protect the privacy of voice data.

The constrained resources of the mobile VoIP devices raise the challenges to design an efficient key updating with voice quality preserving. In order to avoid slowdown from key updating, time-consuming operation should be eliminated. How to realize the secure key updating while preserving the speech quality in the VoIP voice transmission? In this study, we propose a dynamic key updating scheme by suing AES and steganography technologies. Different from traditional cryptology based solutions, we adopt steganography technology to resist the attacks associated with shared keys in VoIP environments. In the proposed scheme, the key materials encrypted by AES are embedded in some VoIP voice packets by using the steganography algorithm. Since the encrypted key materials are hid in some VoIP packets, the attackers cannot obtain the new shared key materials even if the adversary compromises the encrypted key. In order to preserve the quality of VoIP calls, only two high entropy random integers encrypted by AES need to be exchanged in the proposed scheme. Moreover, the two encrypted random integers are embedded in some VoIP packets, so no extra VoIP packets are needed to transmit during the key updating process. In this solution, the quality of the VoIP call can be guaranteed between the end users by employing lightweight AES and steganography algorithm. Therefore the proposed scheme realizes the secure key updating without affecting the quality of the VoIP calls.

Next, we present the proposed dynamic key updating scheme in detail. We assume that two communication parties possess a secret shared key generated in previous authenticated key agreement process during the VoIP call setup phase. In order to protect the transmitted key materials in an unsecure public channel, Advance Encryption Standard (AES) encryption algorithm is adopted to encrypt the key materials. And steganography is employed to hide the existence of encrypted key materials to reduce the probability of attacks efficiently. The dynamic key updating scheme during voice communication is described in detail as follows.

Step 1. After a time period of VoIP calls, when shared session key update is requested, user Alice delivers an INVITE message to Bob via SIP server to start the key updating process.

Step 2. If Bob is busy, "183 Session in Progress" message as a response will be sent back to Alice to acknowledge its busy status.

Step 3. Upon receiving the "183 Session in Progress" message, Alice chooses a high entropy random integer a and then encrypt it by using Advance Encryption Standard (AES) encryption algorithm. In the first key updating process, the shared key negotiated from the authenticated key agreement process is adopted as the encryption key. After that, the new shared key generated from the key updating process will replace the old shared key, and the new shared key will be used as an encrypted key.

Step 4. Alice adopt MD5 algorithm via the shared key K to generate a hash-value $h(K)$, which is used to determine the final embedding location in VoIP packet. Next Alice extract

an integer r from the hash-value $h(K)$, and then she uses the integer r as a random seed to initialize the srand function for generating a sequence of pseudo-random numbers by constantly calling the function of rand. Consequently, Alice embeds the encrypted random integer $E_K(a)$ in a VoIP packet. To complete the hiding operation, the bits stream of the encrypted value substitute the least significant bits of speech sample which determined by the pseudo-random sequence. Then, the encrypted key materials can be transmitted to the Bob unobserved.

Step 5. After receiving the VoIP packet, the end user Bob extracts the hidden data $E_K(a)$ and decrypts it to get the high entropy random integer a by using the shared key. Next, similar to *Step* 3 and 4, Bob selects a high entropy random integer b and adopts AES encryption algorithm to encrypt it with a shared key. Consequently, this encrypted random integer is hid in a VoIP packet and transmitted to the Alice securely.

Step 6. When Alice receives the VoIP packet with encrypted key materials, she first extracts the hidden data $E_K(b)$ from the VoIP packet and then decrypts this data to obtain the high encrypt random integer b. Next, Alice computes the new shared key $K = a \oplus b$ and transmits the "200 OK" message to Bob.

Step 7. After receiving the"200 OK" message, Bob computes the new shared key $K = a \oplus b$ and replaces the previous shared key with the new one. Finally, he sends the "200 OK" message to Alice.

Step 8. When Alice receives the "200 OK" message, she replaces the previous shared key with the new one $K = a \oplus b$.

III. SECURE ANALYSIS

The security of the proposed key updating scheme is based on the security of AES and steganography algorithm used for encrypting and transmitting the key materials. Since AES encryption mechanism as a secure and efficient symmetric encryption algorithm has been applied to encrypt end-to-end voice traffic by Skype. In this section, the security analysis is focus on steganography algorithm which is used to hide existence of the delivered key materials. The Mann-Whitney- Wilcoxon (M-W-W) test [12] is one of the best statistical analysis approaches. Compared with traditional statistical analysis, it has a wider range of applications due to its advantages such as simple computation, rapid speed and time saving. So, we employ M-W-W test to evaluate the security of our scheme.

The Mann-Whitney-Wilcoxon test is a non-parametric test which is used to evaluate whether the stego VoIP stream and the normal VoIP stream as observed independent samples are indistinguishable by comparing the probability distributions between them. In this study, M-W-W test was employed to evaluate the security of the proposed steganographic algorithm. The M-W-W test is based on the standardized test statistic:

$$z^* = \frac{S_2 - E\{S_2\}}{\sigma\{S_2\}}$$

Where $E\{S_2\}$ and $\sigma\{S_2\}$ denote the mean and square root of variance of the sampling distribution S_2, respectively. According to statistical analysis, to have 95% confidence, i.e., with a confidence coefficient (1-α) of 0.95, where α is called the level of significance, we therefore require z (1-α/2) = z (0.975) = 1.960, where z is the percentile of the standard normal distribution. Hence, the decision rule for the test is as follows:

If $|z^*| \le 1.960$, conclude H_0 (two distributions do not differ);

If $|z^*| > 1.960$, conclude H_1 (two distributions differ).

Table 1 shows that the M-W-W test results of comparing the probability distribution drawn from the original VoIP streams to that drawn from the stego VoIP streams with the encrypted key materials. Since $|z^*| \le 1.960$, we concluded H_0–that the probability distributions for the stego-speech with the encrypted key materials and the original cover-speech without embedding did not differ. This means the encrypted key materials hidden in VoIP packets with the steganographic algorithm is undetectable by using statistical analysis.

TABLE 1. M-W-W TEST RESULTS

Steganography	z^*	H
Female 1	0.0483	H_0
Female 2	-0.0417	H_0
Male 1	-0.0875	H_0
Male 2	0.1439	H_0

IV. EXPERIMENTS

To evaluate the performance of the proposed steganographic algorithm, we employed the VoIP speech samples coded by PCM as the cover-speech. The speech samples employed were classified into two groups, female speech samples and male speech samples. We implemented two sets of tests upon the female speech samples and the male speech samples, respectively. To evaluate the impact of steganography upon speech quality, we measured Perceptual Evaluation of Speech Quality (PESQ) scores of the stego-speech with steganography. In PESQ measurements, the cover-speech played by the caller was used as the reference signal, and the stego-speech was measured as the degraded signals.

TABLE 2. TESTING RESULTS FOR FEMALE AND MALE SPEECH SAMPLES

Stego-speech samples	PESQ		
	Max	Min	Mean
Female 1	4.44	4.41	4.425
Female 2	4.54	4.54	4.540
Male 1	4.34	4.30	4.320
Male 2	4.49	4.49	4.490

Table 2 lists the average PESQ values of female and male stego-speech samples with the encrypted key materials. The PESQ values of stego-speech samples changed little, and the PESQ values were 4.4 approximately. It is obvious that the proposed encryption and embedding operations had caused

little degradation in speech quality. Therefore, the proposed scheme achieved covert VoIP communication effectively.

V. CONCLUSION

In this study, we proposed a dynamic key updating scheme without affecting the quality of VoIP call. Unlike other works, both steganography technology and encryption were employed to achieve secure and efficient dynamic key updating during a VoIP call. In the proposed scheme, AES algorithm was used to transform the key material into cipher text to obfuscate the transmitted data and steganography technology was employed to hide the existence of the encrypted key materials to obfuscate the fact of communication. Therefore, the proposed scheme provided protection of key updating to resist various possible attacks efficiently. Furthermore, since the encrypted key materials were embedded in some VoIP packets during the key updating process, the proposed scheme alleviated the burden of network bandwidth and preserved the quality of the VoIP call. The security analysis demonstrated that the proposed scheme enhanced the security of key updating for VoIP by applying steganography with AES algorithm. Therefore the proposed scheme realized secure key updating without degrading the quality of VoIP call in an unsecure Internet environment.

REFERENCES

[1] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler: SIP: Session Initiation Protocol. Internet Engineering Task Force, RFC 3261(2002).

[2] M. Handley and V. Jacobson: SDP: Session Description Protocol. RFC 2327(1998).

[3] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson: RTP: A Transport Protocol for Real-Time Applications. Internet Engineering Task Force, RFC 3550(2003).

[4] A. Lahmadi and O. Festor: A framework for automated exploit prevention from known vulnerabilities in voice over IP services. IEEE Transactions on networks and service management, 9, 2,114-127 (2012).

[5] Chia-Hui Wang, Yu-Shun Liu: A dependable privacy protection for end-to-end VoIP via Elliptic-Curve Diffie-Hellman and dynamic key changes. Journal of Networks and Computer Applications, 34, 5, 1545-1556 (2011).

[6] H. Hakan, Kilinc, and Tugrul Yanik: A Survey of SIP Authentication and key Agreement Schemes. IEEE communications Surveys and Tutorials, 16, 2, 1005-1023(2014).

[7] D. Wang, D. He, P. Wang, C. Chu: Anonymous Two-Factor Authentication in Distributed Systems: Certain Goals Are Beyond Attainment. IEEE Transactions on Dependable and Secure Computing, DOI: 10.1109/TDSC.2014. 2355850, 2014.

[8] H. Arshad, M. Nikooghadam: An efficient and secure authentication and key agreement scheme for session initiation protocol using ECC. Multimedia Tools and Applications, DOI: 10.1007/s11042 -014-2282-x, 2014.

[9] T. Hand, K. Neeraj, C. Naveen, R. Resungmin: An improved authentication protocol for session initiation protocol using smart card. Peer-to-Peer Networks and Applications, DOI: 10.1007/s12083-014-0248-4(2013).

[10] M. Farash, M. Attari: An anonymous and untraceable password-based authentication scheme for session initiation protocol using smart cards. International journal of communication systems, DOI 1002/dac.2848 (2014).

[11] B. Feng, W. Lu, W. Sun: Secure binary image steganography based on minimizing the distortion on the texture. IEEE Transactions on Information Forensics and Security, 1, 2, 243-255 (2015).

[12] Neter J, Wasserman W, Whitmore G A.: Applied Statistics. 4th ed. Simon & Schuster, 435-450(1993).