Extending AES with DH Key-Exchange to Enhance VoIP Encryption in Mobile Networks

Raid Zaghal(✉), Saeed Salah, and Noor Jabali

Department of Computer Science, Al-Quds University, Jerusalem 20002 Israel
{zaghal, sasalah}@staff.alquds.edu,
Noor.jabali@students.alquds.edu

Abstract. Due to the huge developments in mobile and smartphone technologies in recent years, more attention is given to voice data transmission such as VoIP (Voice over IP) technologies, *e.g.,* (WhatsApp, Skype, and Face Book Messenger). When using VoIP services over smartphones, there are always security and privacy concerns like the eavesdropping of calls between the communicating parties. Therefore, there is a pressing need to address these risks by enhancing the security level and encryption methods. In this work, we suggest a new scheme to encrypt VoIP channels using (128, 192 & 256-bit) enhanced encryption based on the Advanced Encryption Standard (AES) algorithm, by extending it with the well-known Diffie-Hellman (DH) key exchange method. We have performed a series of real tests on the enhanced (AES/DH) algorithm and compared its performance with the generic AES algorithm. The results have shown that we can get a significant increase in the encryption strength at a very small overhead between 4% and 7% of execution time.

Keywords: VoIP · AES · Diffie-Hellman · Key exchange · Security Encryption

1 Introduction

In recent years, smartphones have become an integral part of almost all people's lives around the world. The huge developments in mobile applications have made our lives easier, yet more sophisticated and more challenging (*e.g.,* due to security concerns). It is no longer just a phone with a SIM card for voice calls only, but it is now being used in many ways due to the rapid developments in hardware and software in mobile phone technology. For example, mobile phones (smartphones) use has expanded to include sending audio and video messages, checking emails, shopping, banking services, browsing the Web, accessing the cloud services, navigating maps, camera services, and many more applications. One of the mostly used and useful technologies in mobile phones is the live streaming of media including voice communications known as Voice over Internet Protocol (VoIP).

VoIP in general is a demanding, cost ineffective technology that entails semi-free voice calls carried through the Internet, and therefore it draws the attention of many Internet users. However, VoIP deals with small-size packets compared with other video

Á. Rocha et al. (Eds.): WorldCIST'18 2018, AISC 746, pp. 511–521, 2018.
https://doi.org/10.1007/978-3-319-77712-2_48

or Web packets, and the redundant header size of a VoIP packets can be larger than the size of the payload including voice information [1]. Therefore, this kind of voice communication is vulnerable to diverse types of attacks such as eavesdropping, and packet capturing, among others. To protect this communication channel from these attacks, users need to employ certain protective measures such as authentication, confidentiality (via encryption), and integrity with replay protection to the media stream. The most important protective measure is confidentiality of the data which means that the encrypted data is indistinguishable by anyone who does not have the key. Message authentication implies that if a second party user agent (a server) receives a packet sent by a first party user agent (a client), then it is indeed sent by the first party. Data integrity implies that any modification of the data during transmission will be detected by the recipient [2].

Despite the substantial number of studies that proposed techniques to handle the security issue in VoIP service [3–8], this topic is still an active area of research. This is due to the fact that the deployment of VoIP technologies across the Internet encounter several challenges such as architecture complexity, QoS issues, and security concerns. In addition, security problems are becoming more complex and thus traditional security equipment, protocols, and architectures cannot adequately behave very well against recent sophisticated and intelligent attacks like traffic analysis and social engineering/semantic attacks. Currently, VoIP uses several ways to protect data, such as different algorithms to encrypt data packages like Advanced Encryption Standard (AES) [2], Data Encryption Standard (DES) [9], Diffie-Hellman (DH) [10], Rivest Cipher (RC2) [11], Blowfish (BF) [12], and Triple DES, among others. But, recently with the introduction of new methods of decryption these algorithms start to have problems in data protection in VoIP. For example, the AES algorithm has a problem with the level of security by using weak key exchange; there are many cryptanalysis systems that were designed to attack this algorithm using the same key to decrypt the packet or attack block ciphers with a reduced number of rounds. Therefore, its necessary to enhance the level of security by using a strong asymmetric key exchange, and making it almost impossible to hack by different intruders or against any kind of attacks such as differential, brute-force and liner attacks.

The main contribution of this paper is to extend the AES encryption algorithm (which is known for its high speed and reasonable security) with the DH key-exchange algorithm (which is characterized by its strong security and reasonable speed), with the aim of reducing the threats during voice transfer operation by increasing the security level and making it almost impossible for intruders to compromise the traffic channel. As the results showed, by incorporating these two algorithms in one process, we gain a significant added-value with a very small overhead. We call this new extension (AES with DH extension) and compare it with the classical (AES) approach, as will be shown in the experimental section.

The remaining of this paper is structured as follows. First, a review of the related work is presented in Sect. 2. Then, the extended version of the AES algorithm with DH encryption process is discussed in more detail in Sect. 3. Section 4 presents the validation analysis of the method. The proposed method and its applications are experimentally tested and evaluated in Sect. 5. Finally, in Sect. 6, we draw various conclusions and provide some insights into further works.

2 Related Work

During the last few decades, there were huge developments in communication technologies and specially in mobile technology. The biggest leaps were made in the Internet related technologies and services [10]. These developments have caused many emerging problems related to security and privacy, *e.g.,* an intruder can discover the essential information or data through reading the text or eavesdropping on the voice channel, the delay incurred through the transfer process can increase to undesirable levels, and the degradation of session reliability can cause side-effects on the performance of the system. Thus, the voice transmission process from the sender to the receiver should provide adequate protection via a secure/encrypted channel.

To solve these issues, during the past years, several techniques and algorithms have been proposed to improve the quality of VoIP transmission, but many of these techniques have focused on one side of the transmission aspect, such as security, quality or speed. In what follows we list the most related work and summarize their main differences with our work.

Prashant and Trimbak [13] proposed a Peer-to-Peer (P2P) media application called SALMON, to secure video streaming utilizing the SIP proposal. It was built against meddlers between pairs that use some media to exchange data. The application is used by elliptic-curve Diffie-Hellman (EDH) key trade on SIP flagging and AES encryption. SALMON uses Omnet++ for building applications to capture quality of control messages, disruption count and service latencies. The results have shown that the average transfer speed and life time are associated, not just have a finer execution in overhead of control messages.

Rebahi *et al.* [14] presented a Request for Comment (RFC 4474) standard for the SIP protocol used in managing IP multimedia sessions in the Internet. One of the main conclusion of this report is that it strongly recommends the use of the DH algorithm for securing multimedia sessions over the Internet. They claim that the proliferation of small and simple devices as well as the need to increase the capacity of the SIP servers to handle the increasing VoIP traffic will make continuous reliance on DH more challenging over time.

Vargic, *et al*. [15] developed a novel service architecture based on second Generation (2G) subscriber that accesses a SIP based VoIP network via Wi-Fi complying security standards, such an approach can be used especially in highly-populated areas, such as airports and business centers. This architecture suffers from several issues: First, that IP Multimedia Subsystem (IMS) suffers from lack of clients. Second, the mobile operators want to give the subscriber a possibility to access their VoIP network and efficiently cover densely populated areas. User authentication and authorization are based on an algorithm that uses the Extensible Authentication Protocol (EAP) and proves the user identity by owning a Subscriber Identity Module (SIM). This algorithm is called the EAP-SIM algorithm. The integrity and confidentiality are provided by Internet Protocol Security (IPSec) connection established using parameters derived from authentication triplets. They tested secured and non-secured SIP sessions. The latter ones were tested for non-IPSec enabled clients, and verified the proposed architecture with a mobile phone, and have proven the correctness of their approach.

The main drawback that remains is the difficulty of IPSec implementation that can be passed by a special application. The lack of IMS clients and specific requirements of the mobile operators have forced researchers and industry to develop new service architectures. Son *et al.* [3] proposed an encryption model for the VoIP service using 128-bit AES to reduce the voice delay of the VoIP telephone. The call process uses a mutual key exchange in the encryption system and asymmetric encoding method RSA algorithm to improve security. The results have shown a satisfactory performance in terms of Mean Opinion Score (MOS) and an R-factor.

Kilinc and Yanik [16] presented a survey of authentication and key agreement protocols that are critical security services to implement a secure SIP protocol which is a common part of the VoIP architecture. They concluded that performance and security of the authentication and key agreement schemes are two critical factors that affect the VoIP applications with substantial number of users. Therefore, the performance of the authentication and key agreement protocols is significant. Furthermore, they identified, categorized and evaluated various SIP authentication and key agreement protocols according to their performance and security evaluation. They compared the different schemes according to four various categories they are: Password Authenticated Key Exchange (PAKE), Hash functions, Public Key Cryptography (PKC) based and ID based schemes. Singh *et al.* [17] discussed many issues that the Internet has revolutionized the telecommunication systems by supporting new applications and services. Voice over Internet Protocol (VoIP) is one of the most prominent telecommunication services based on the Internet Protocol (IP). The signal quality of the VoIP system depends on several factors such as networking conditions, coding processes, speech content and error correction schemes. From the very beginning of transferring the voice data over packet switched networks, the journey of the packet based communications to modern VoIP and advancements to improve the service of the VoIP system. The VoIP system has been established as the best alternative to the traditional Public Switched Telephone Network (PSTN) telecommunication system for providing the voice services to the users.

In summary, most of the previously listed contributions do not consider delays introduced by the network and database access in the design. It is worth noting that when designing authentication and key agreement protocols it is necessary to consider the delays, especially in distributed computing environments. For P2P and Next Generation Networks (NGN) architectures, new authentication and key agreement protocols that consider the various overheads introduced by the distributed network structure are necessary. Thus, the work presented in this paper differs in such a way that it attempts to handle several important aspects of the VoIP channel by incorporating a SIP server to monitor the communication session between the sender and the receiver, and most importantly, it proposes an extension of the AES encryption algorithm by incorporating the DH key exchange algorithm to enhance the security level of the voice channel. It is worth mentioning that in this work we have tested the efficiency of the proposed method against brute force attacks that are widely used by attackers through application programs to attempt to decode the encrypted channel rather than employing systematic methods like traffic analysis and social engineering.

```
00: initialize all variables
01: if (in-call == New Call)
02:  Connect with SIP server
03:  Generate (10 Keys) from IP-Sender.
04:     Send required extension to the SIP
05:  Generate 10 Keys from IP-receiver.
06:  Start the connection
07:     K ← Chose Random (Key)
08:     K1 ← encrypt by DH (K)
09:  Send K1 to the receiver.
10:  if (Flag == False)
11:   Flag ←True
12:  else
13:   Flag ← False
14:  end if
15:  if (Flag ==True)
16:   J ←2
17:  else
18:   J ←1
19:  end if
20: end if
21: if (in-call == False)
22:  Record from MIC by Stream Thread
23: end if
24: while (in-call == True)
25:  D ← 8096 byte (data)
26:  for I=J: 2:10
27:   D1 ← Encrypt data by AES [D]
28:   D2 ← Encrypt by AES [index key]
29:   Send (D1,D2) Send Sound to SIP port
            (50505)
30:   if (in-call== True)
31:    Continue
32:   else
33:    Break
34:   end if
35:  end For
36: end while
37: Return (D1, D2)
38: Output: encrypted audio
```

Listing 1. The modified version of the AES encryption algorithm with DH extension.

3 AES with DH Extension

We incorporate the idea of the Diffie-Hellman algorithm into the AES process. The extended version, *i.e.* AES with DH, is implemented by adding additional new interface between the two parties. It is built outside the SIP server.

The sender part uses the AES/DH extended algorithm for encrypting voice before transmitting it to the other party by the SIP server, and the receiver will use the same algorithm in reverse to decrypt the voice. Listing 1 describes the pseudo code of the extended version of the AES algorithm with the DH key exchange (The encryption part at the Sender), while Listing 2 describes the decryption part at the receiver.

```
00: initialize all variables
01: X← Received (K1) from sender
02: X1← DH(X)
03: while (in-call == True)
04:  Received (D1) and (D2) from port (50505)
05:  X2← D1
06:  X3← decrypt key index by AES [key index]
07:  X4← decrypt data by AES [X2]
08:  X5← combine (X3,X4)
09:  Listen (X5)
10: Return (X5)
11: Output: original audio
```

Listing 2. The modified version of the AES decryption algorithm with DH extension.

Listing 1 shows the steps of our encryption strategy (AES/DH). This process is applied at the sender and opens the session with the SIP server; it also checks if the second party is available to receive data. After this initial setup, the sender starts the encryption process as follows:

- **Start call:** When the sender starts a call, the first step is to open a session with SIP server to manage the call and make sure that the second party is available for connection.
- **Exchange key:** Generate 10 keys from both users to use them in the DH algorithm to encrypt any random key, and to start the key exchange between the two parties and send the desired extension to the SIP server, and thus secure the channel.
- **Encrypt data:** The sender uses the keys generated by the DH algorithm to encrypt the data, indexed by AES algorithm, and sends it to the second party by the Audio Streamer – sound to port (50505).
- **End call:** Continue the exchange of data (audio) until one of the parties ends the call.

Listing 2 presents the decryption steps applied by the receiver to decrypt the data according to the following:

- **Receive key:** Receive the DH key from the sender.
- **Receive data:** Receive the data and index from sender by audio receiver from port (50505) that will be used by the AES decryption algorithm.

- **Decrypt index:** After receiving the index, the AES algorithm is used to decrypt the key.
- **Decrypt data:** After receiving the data, the AES algorithm used to decrypt it.
- **Combine results:** Combine the results that we got from the index and data to generate the original file and listen to voice.

4 Validation

Here we provide some analysis on the validity of our technique and its worth-fullness versus the small amount of overhead incurred by the added work of employing the DH algorithm on top of the AES. In this module we assume that both algorithms were attacked by brute force attackers for different key sizes (128, 192 and 256-bit keys), and we show the possible number of key combinations. The number of combinations for each algorithm will take from rounds to check every possible key combination starting with "0000." Given sufficient time, a brute force attack can crack any known algorithm. For example, the brute force attack on a 4-bit key is will take a maximum 16 rounds to check every possible key combination. Table 1 summarizes these numbers.

Table 1. Possible number of key combinations.

Key size	Algorithm	
	AES combinations	AES/DH combinations
128 bit	$3.4x10^{19}$	$11.56x10^{76}$
192 bit	$6.2x10^{28}$	$38.44x10^{114}$
256 bit	$1.1x10^{38}$	$1.21x10^{154}$

Table 2 shows the probability of the hacker's success to decipher the channel for both AES and AES/DH. It is obvious that hacker's chances to succeed are much harder with the AES/DH method, it will take him/her many billions of years to manage to get the correct key combinations; and this validates our original claim that we can reach a total new level of security with this added layer and we have made it almost impossible for any hacker to compromise the voice channel and decrypt the VoIP data.

Table 2. The probability of the hacker's success to decipher the channel.

Key size	Algorithm	
	AES probability	AES/DH probability
128 bit	0.980×10^{-18}	2.87×10^{-57}
192 bit	0.534×10^{-37}	8.62×10^{-96}
256 bit	0.302×10^{-56}	2.74×10^{-134}

5 Experimental Results

In this section we summarize the experimental results and list the main findings. We run the experiments using six scenarios and in each scenario, we repeated the same experiment 20 times with a total of 120 runs. The six scenarios and their descriptions are listed in Table 3.

Table 3. The six implemented scenarios and their descriptions.

Scenario#	Parameter			Algorithm	
	Key size (bit)	Packet size (Byte)	Router (Mb/s)	AES	AES&DH
1	128	20000	108	√	√
2	192	20000		√	√
3	256	20000		√	√
4	256	10000		√	√
5	256	40000		√	√
6	256	10000	54	√	√

The implementation was run on a PC with CPU speed 1.70 GHz and 8 GB RAM. We used Samsung mobile device (Galaxy S5) with CPU speed of 1 GHz and 8 GB RAM. The network speed is chosen to be either 54 Mbps or 108 Mbps. To create the six scenarios, we have used the Android Platform Development (Android Studio), and tested the application on real devices with a SIP server (3CX). The CX Phone System is a SIP server under Windows OS that works with popular VoIP Gateways; SIP phones allow you to setup a complete IP, and the SIP servers are responsible for setting up the calls between SIP devices and usually combine several functions including SIP proxy and SIP registers into one piece of software. We also used the *Wireshark Capture Filter* [18] – which is a network protocol analyzer – to capture the traffic while running the experiments. It is a free software that supports many protocols and media such as real data from 802.11 Wireless LAN, Ethernet, Token-Ring, FDDI, and ATM connections. Finally, we used two mobile devices running the Android 4.1 (JellyBean) Operating System.

To study the overhead induced by incorporating the DH key exchange idea into the AES process, we captured three parameters while running the sessions: (1) execution time for both encryption and decryption processes, (2) propagation delay, and (3) number of lost packets. We repeated the experiments twice: first time we engaged the AES algorithm only, and the second time we engaged our version (the AES with DH combined).

Figure 1 presents the experimental results of the three parameters. We fixed the packet size to 20 KB and studied the effect of increasing the key size (128, 192, and 256-bit) to the performance of both algorithms (AES vs. AES/DH). Figure (1.A) gives the average encryption time, Fig. (1.B) gives the average decryption time, Fig. (1.C) presents the average propagation delay, and finally, Fig. (1.D) gives the number of lost packets. It is obvious that in the four plots that the AES algorithm is slightly faster than the AES/DH. For example, if we average the encryption time for the three cases

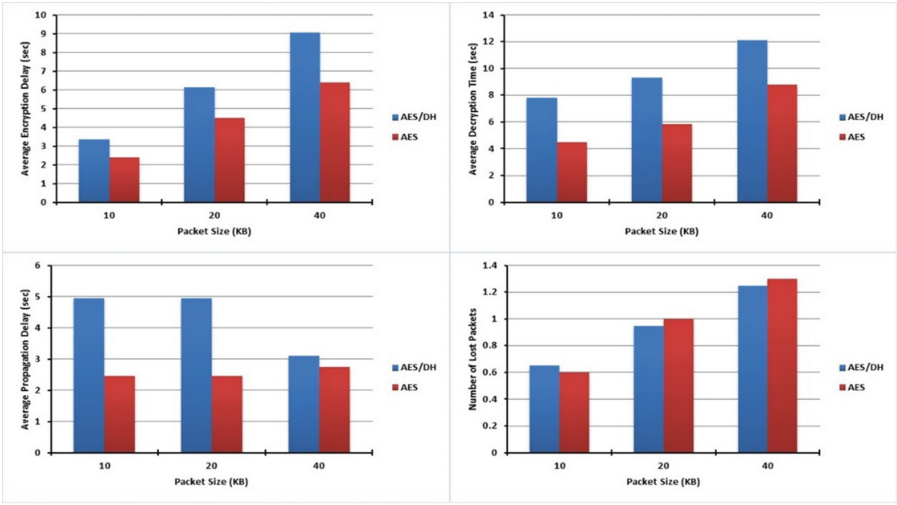

Fig. 1. Effect of increasing key size to Enc./Dec. times, progapation delay, and packet loss.

(key sizes) we get 3.8 ms for the AES vs. 5.5 ms for AES/DH alone; or 1.7 ms increase. In the same way, if we average the decryption time, we get 24 ms for the AES/DH vs. 16.5 ms for AES alone; around 8 ms increase. If we repeat the same analysis for the propagation delay we get around 4.5 ms increase, and on average the number of lost packets is almost the same for both cases. It is obvious that we have added some overhead due to the extra work being done specially during encryption and decryption processes to engage the DH algorithm, but we can say that the summation of all three times (encryption + decryption + propagation delay) has increased by 14.5 ms in total (average of the three key sizes), which is totally acceptable for voice communication.

To get more insights into the behavior of both algorithms, we carried another set of experiments by fixing the key size to 256-bit and tested with three packet sizes (10 KB, 20 KB, and 40 KB). The experimental results of this test are shown in Fig. 2. Here we used the same four plots as in the previous figure. Based on this set of experiments we find: *(i)* there is an added overhead in the three-time plots (encryption+decryption+propagation delay), but for the number of lost packet, the performance is equal, *(ii)* It is also obvious in these plots, that the larger the packet, the better (*i.e.,* less overhead is added). *(iii)* if we perform the time analysis by averaging the total times for the three packet sizes, we get the following numbers: encryption time has increased by 5 ms, decryption time has increased by 10 ms, and propagation delay time has increased by 5 ms. Therefore, the average total overhead is around 20 ms, which is compatible with the pervious result (14.5 ms).

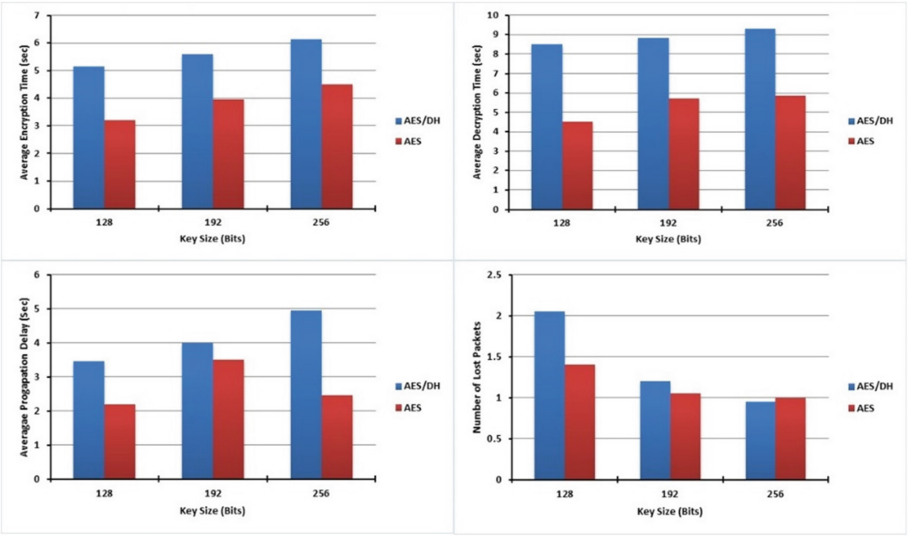

Fig. 2. Effect of increasing packet size to Enc./Dec. time, propagation delay, and packet loss

6 Conclusion and Future Work

In this work, we have designed a new strategy for securing a VoIP channel against brute force attacks by combining the AES encryption algorithm and the DH key-exchange algorithm. We have experimented with a real testbed including two mobile devices and a SIP server. We have shown that the new strategy will increase the security level of the voice channel at a very small overhead. Since VoIP calls are usually low-cost or semi-free service these days, it is currently wide spread among billions of Internet users; and therefore, VoIP calls can be subject to several types of attacks like eavesdropping, capturing packets, and interference. So, it is of great importance that we increase the security of these channels to maintain high level of privacy, confidentiality and reliability for the users. By combining the two algorithms in one process that we have called (AES/DH); we have benefited from AES's speed and DH's strength. In the future, we plan to investigate and experiment our strategy with other algorithms and other kinds of media like video and chatting, and probably experiment on a bigger testbed. Besides, we plan to modify the proposed strategy to consider other types of attacks such as traffic analysis and social engineering methods.

References

1. Jung, J.Y., Kang, H.S., Lee, J.R.: Performance evaluation of packet aggregation scheme for VoIP service in wireless multi-hop network. J. Ad Hoc Netw. **11**(3), 1037–1045 (2013)
2. Daemen, J., Rijmen, V.: Advanced Encryption Standard (AES). Published by NIST: FIPS PUB 197 (2001)

3. Son, B., Nahm, E., Kim, H.: VoIP encryption module for securing privacy. Int. J. Multimedia Tools Appl. **63**(1), 181–193 (2013)
4. Munef, Z.: Securing VoIP in SIP mobile network. J. World Comput. Sci. Inf. Technol. **5**(1), 6–10 (2015)
5. Shan, L., Jiang, N.: Research on security mechanisms of SIP-based VoIP system. InL Proceedings of the 9th International Conference on Hybrid Intelligent Systems (HIS 2009), August 12–14, 2009, Shenyang, China (2009)
6. Pradhan, C., Bisoi, A.K.: Chaotic variations of AES algorithm. Int. J. Chaos Control Model. Simul. **2**(2), 19–25 (2013)
7. Thiruppathi, M.S.: Improving quality in Voice Over Internet Protocol (VOIP) on mobile devices in pervasive environment. J. Comput. Appl. (EICA 2012-4), 334–339 (2012)
8. Mark, C.: Basic vulnerability issues for SIP security. Chief Technology Officer Secure Logix Corporation, March 2005
9. Karthik, S., Muruganandam, A.: Data encryption and decryption by using triple DES and performance analysis of crypto system. Int. J. Sci. Eng. Res. IJSER **2**(11), 24–31 (2014)
10. Diffie, W., Hellman, M.: New directions in cryptography. IEEE Trans. Inf. Theory **22**(6), 644–654 (1976)
11. William, J.: RC2 Encryption in .NET (2011). http://buchananweb.co.uk/security06.aspx
12. NehaKhatri, V., Kshirsagar, V.: Blowfish algorithm. IOSR J. Comput. Eng. **16**(2), 80–83 (2014)
13. Kumbharkar, P.B.: A secure overlay dynamic multicast network with load balancing for scalable P2P video streaming services. Int. J. Adv. Comput. Technol. **3**(10) (2014)
14. Rebahi, Y., Minh, N.T., Zhang, G.: Performance analysis of identity management in the Session Initiation Protocol (SIP). In: Proceedings of the IEEE/ACS International Conference on Computer Systems and Applications (AICCSA) (2008)
15. Vargic, R., Kotuliak, I., Vrábel, A., Husák, F.: Provisioning of VoIP services for mobile subscribers using Wi-Fi access network. Telecommun. Syst. **52**(3), 1705–1711 (2013)
16. Kilinc, H.H., Yanik, T.: A survey of SIP authentication and key agreement schemes. IEEE J. Commun. Surv. Tutorials **16**(2), 1005–10230 (2014)
17. Subashri, T.: Confidentiality of VoIP data using efficient ECDH key exchanging mechanism. In: Proceedings of the 8th International Conference on Applied Electromagnetics, Wireless, and Optical Communications (2014)
18. Wireshark. https://wiki.wireshark.org/CaptureFilters