A Survey of Different SPIT Mitigation Methods and a Presentation of a Comprehensive SPIT Detection Framework

Mohammad Hossein Yaghmaee Moghaddam, Mina Amanian, Farideh Barghi, and Hossein Khosravi Roshkhari

***Abstract*—VoIP is a promising technology for voice transmission on IP-based networks and it has many advantages over PSTN. One of the most important threats in these networks is unsolicited bulk calls, known as SPIT (Spam over Internet Telephony). Our purpose in this paper is doing a deep research into this topic and presenting a new anti-SPIT mechanism.**

In order to detect SPIT efficiently we need to extract some features which help us in categorizing the incoming calls. In this paper we propose an approach based on extraction of important features that contain all aspects of call, hence it can detect SPIT efficiently in an acceptable time. In this paper, eight features which are directly extracted from the SIP header are applied in the detection process.

The simulation results show that the proposed framework is a comprehensive and efficient solution which provides acceptable true-positive and false-negative values.

***Index Terms*—PSTN, SPIT, SIP, VoIP, Caller, Callee.**

I. Introduction

VoIP is a promising technique which uses existing data networks to establish voice sessions via transferring voice streams replaced into data packets. Nowadays voice transmission via internet has become an essential tool for business market which yields in development and efflorescence of it. VoIP is a transformation of traditional networks or PSTNs (Public Switched Telephone Networks) proposed in Data Network environment. The main reason of VoIP fast development is that it reduces telephony costs, produces higher availability and provides easy convergence to PSTN. We must perform mechanisms to support real-time delivery for voice packets and also signaling protocols for handling the communicational negotiations between different VoIP devices.

Threats and vulnerabilities of internet protocols make VoIP potentially insecure, for example an attacker can send numerous simultaneous advertisement calls or messages to other users with low cost. SIP (Session Initiation Protocol) is the most applied protocol in current VoIP implementations. SIP is a signaling protocol for initiating, managing and terminating voice and video sessions across packet networks. SIP sessions involve one or more participants and can use for unicast or multicast communication [1].

Attackers attempt to establish bulk unsolicited multimedia sessions which are called Spam calls [2]. This type of Spam calls in VoIP environment is named SPIT (Spam Over internet telephony). Although SPIT is essentially similar to Spam but we expect that SPIT has more destructive effect on user [2], [3]. Due to the expensive and end to end nature of PSTN, generating Spam calls are less attractive to attackers in such networks [4].

Basically attackers use specific software (which is assumed as soft phone client) named bot. The attacker may set up Virus or Trojan on various computers to distribute the Caller identity. First report took place in 2006 that many advertisement calls with marketing purpose showed up in Skype. Spam detection approaches are not extendable to SPIT due to the real time nature of it. Moreover, Email content plays a big role in detection mechanism which is not applicable for VoIP content [4]. The similarity of SPIT and Spam is that both of them use internet to achieve their purposes. Due to the wide deployment of VoIP networks, we assume that SPIT will be a serious problem in the near future.

This paper is organized as follows. In Section II we survey the related works on Anti-SPIT methods in details. Section III describes the background, section IV discusses the proposed mechanism, Section V provides the simulation results, Section VI contains evaluation and finally the conclusion is presented.

II. Related Work

There are different methods to avoid SPIT. These methods are classified in the following groups:

List-based Filtering: tries to detect SPIT by checking specific lists. Callers are categorized within white, black, gray lists. These types of frameworks allow calls which a correspondent equivalent record to callee exists in the white list; blocks or takes other custom actions, if the Caller exists in black list or in gray list respectively. A simple and effective mechanism based on this approach is proposed in [5], when a Callee receives a SPIT call, Caller should press so called SPIT button on the phone. Pressing the SPIT button has two purposes: Add this Caller to the black list and reporting to local proxy servers or public proxy servers. This information is shared within networks and with a simple rule; if the

Manuscript received August 19, 2013; revised December 27, 2013.

F. Mohammad Hossein Yaghmaee Moghaddam is with the Dept. of engineering Ferdowsi University, Mashhad, Iran (e-mail: yaghmaee@ieee.org).

S. Mina Amanian is with the Dept. of Computer Engineering Imam Reza International University, Mashhad, Iran (e-mail:minaamanian@gmail.com, farideh.1367@gmail.com).

F. Hossein Khosravi Roshkhari is with the Dept. of Engineering Ferdowsi University, Mashhad, Iran (e-mail: Hos.khosravi.r@gmail.com).

DOI: 10.7763/IJMLC.2014.V4.410

amount of SPIT feedbacks for a specific user is bigger than a threshold, it will be marked as SPITTER.

Trust-based Filtering: This method uses the body lists and other parameters for scoring the users. It then, allows or blocks calls based on trust or reputation scores [6]. One similar method uses the trust certificates [7]. In proposed mechanism and in similar methods [8], [9], parameters such as call duration are applied in calculating the trust score. For instance, in [7], an average time a Callee usually talks to his white list contacts is calculated. If a new Caller does not acquire a trust value almost close to white list contacts, the new Caller is considered to be suspicious. If the Caller is unknown to the Callee, so is not listed in the Callee white list, the framework uses two solutions; the first is based on the trust propagation. For instance, Alice knows Bob and trusts him, Bob knows Carol too. Therefore, Alice can trust Carol as well. The second is based on a research performed by Microsoft and an analysis on messengers in social networks. They observed that the utmost social distance could be 6.6 between the users. It means 78% of world population can connect to each other within seven steps or less. Therefore, if there wasn't any connection between the Caller and the Callee, it might be marked as SPIT calls and shall be blocked. Special cases, like [10] is proposed for P2P-VoIP.

Interactive-based Filtering: These approaches differ human Callers from automated SPIT generators with Turing tests. This test is based on conversation patterns like silences, response durations and so on. Human voices have special features which could be helpful in SPIT detection. In these techniques, SPIT detection is based on difference between Caller and Callee voice saturation. As usual conversation needs speaking and thinking, therefore the voice saturation ratio might be low. Actually, in normal conversations, voice involves just 40% of whole time and 60% of it is environment noises. In SPIT calls, questions and responses aren't active. A similar approach is presented in [11].

Pattern-based Filtering: By calculating call patterns such as call frequency and comparing it with previous patterns, SPIT calls could be detected [12]-[16]. Suppose three measures. Interaction Rate (IR) means a logic combination of rate of input and output calls from one user. Historical Rate (HR) means repetitive and distinct calls from one user and Social Rate (SR) means the rate of unknown received calls from another user. By combining an optimized value of these three features (X, Y, Z), we can introduce normal behavior. In this mechanism, each new call request must be forwarded from a router and those features should be calculated for it and finally an algorithm is performed to allow or block the call, as shown below:

```
If   ($IR<X$)           then   block
Else If  ($HR<Y$)       then   block
Else If  ($SR<Z$)       then   block
Else                           allow;
```

To summarize, each one of above mentioned methods have advantages and disadvantages. List-based Filtering methods could be implemented very simply. While they show very effective, sybit attacks could expose them to threats. This attack in tries to generate spurious numbers and can change lists falsely.

Interactive-based filtering is appropriate for detecting automated SPITTERs, but it needs processing resources.

Pattern-based filtering and trust-based filtering methods are more flexible and show more effective. The use of trust and reputation provides less false positive and less false negative in the detection process than other methods. False positive means detection of an SPIT as a normal call and false negative means detection of a normal call as SPIT.

Some other approaches use anomaly and ontology for detecting SPIT [3], [12].

Sip Spam labeling system [14] does not use sip extension and uses SIP INVITE message to establish SIP session for the insertion of Spam indicator. Another research collects features of internet telephony [15] then executes k-Nearest neighbor classification and analyses that the user is suspicious or not, this approach immediately updates black list.

While some researchers have focused on the applicability of such solutions taking marketing concepts into consideration [13], it seems that SPIT problem still needs more research.

III. Background

A. SIP

SIP is a peer-to peer protocol with the following entities: User Agents (UA), Proxy Servers, Redirect Servers, Location Servers and Redirect Servers.

In a sample session establishment between two UAs, after registration of the two users, UA_A sends an INVITE request to the Proxy Server. Proxy Server looks up A's IP address and passes the message to UA_B. After UA_B has confirmed the request by phone pick up, UA_A requests a Media session. The established call could be terminated if any of the UAs sends a BYE request to the other UA [1].

SIP is being developed by the SIP Working Group, within the Internet Engineering Task Force (IETF). The protocol is published as IETF RFC 2543 and currently has the status of a proposed standard.

Signaling in SIP is based on (ASCII compatible) text messages. A message is composed of a message header and an optional message body. Messages are either requests or responses. Request messages are sent from the client to the server as shown in Table I and Response messages are sent from the server to the client as shown in Table II [17].

TABLE I: REQUEST MESSAGES SENT FROM THE CLIENT TO THE SERVER

	Method	Description
Request Methods	INVITE	Initiates a call, changes call parameters.
	ACK	Confirms a final response for INVITE.
	BYE	Terminates a call.
	CANCEL	Cancels searches and "ringing"
	OPTIONS	Queries the capabilities of the other side.

Response messages contain numeric response codes. The SIP response code set is partly based on HTTP response codes [18].

TABLE II: RESPONSE MESSAGES SENT FROM THE SERVER TO THE CLIENT

	Method	Description
Response Methods	Register	Registers with the Location Service.
	ACK	Sends mid-session information that does not modify the state.

A SIP-based VOIP comprises of two phases, named, call setup and media session. Call setup establishes a session using a request/response mode. This phase involves in a handshaking between the Caller and the Callee and the media session, to exchange messages between the call participants in an on-going session.

B. SPIT

The term Spam, which is usually used to describe unwanted email, can be expanded to describe any unsolicited message (with positive appearance) in a sample communication. In VoIP networks, this term, which is named SPIT, determines any unsolicited call or message which usually contains commercial data.

Different type of Spam in VoIP networks can be categorized as follows [1]:

- Spam over Internet Telephony (SPIT)
- Spam over Instant Messaging (SPIM)
- Spam over Presence Protocol (SPPP)

SPIT as is generally known, refers to unsolicited calls which usually play pre-recorded audio files for target Callees.

IV. PROPOSED METHOD

Many features can be exploited from a VoIP packet which helps us in SPIT detection. In our proposed scheme we selected eight important features of the call that are easy to implement and fast to extract, they can be exploited directly from SIP header, which are presented as follows:

A. Features

There are many features that could be used for differentiation between the normal calls and the SPIT calls. Among these, call duration and call rate gain the most importance.

- Call duration is the most important feature in all Anti-SPIT mechanisms. Usually call duration in SPIT calls is very short. Because if SPIT calls are established, the Callee will quickly notice it and shall end the communication shortly. But, in normal calls, the communication goes on.

Call duration for each Caller is defined by an average of all call durations during a time period as shown below:

Call duration = Sum of call durations in a period time / number of calls in the same period time

- Call rate: SPITTERs usually send many calls in a certain period of time, while the normal users send few calls in the same period. Therefore, the call rate for SPIT calls is very high. As the purpose of a SPITTER is sending mass calls to many users simultaneously, and to propagate of their advertisement messages, frequent calls in a sample period of time determines the Caller as a SPITTER. This feature is defined as below:

Call rate = number of calls in a period of time / the same period of time

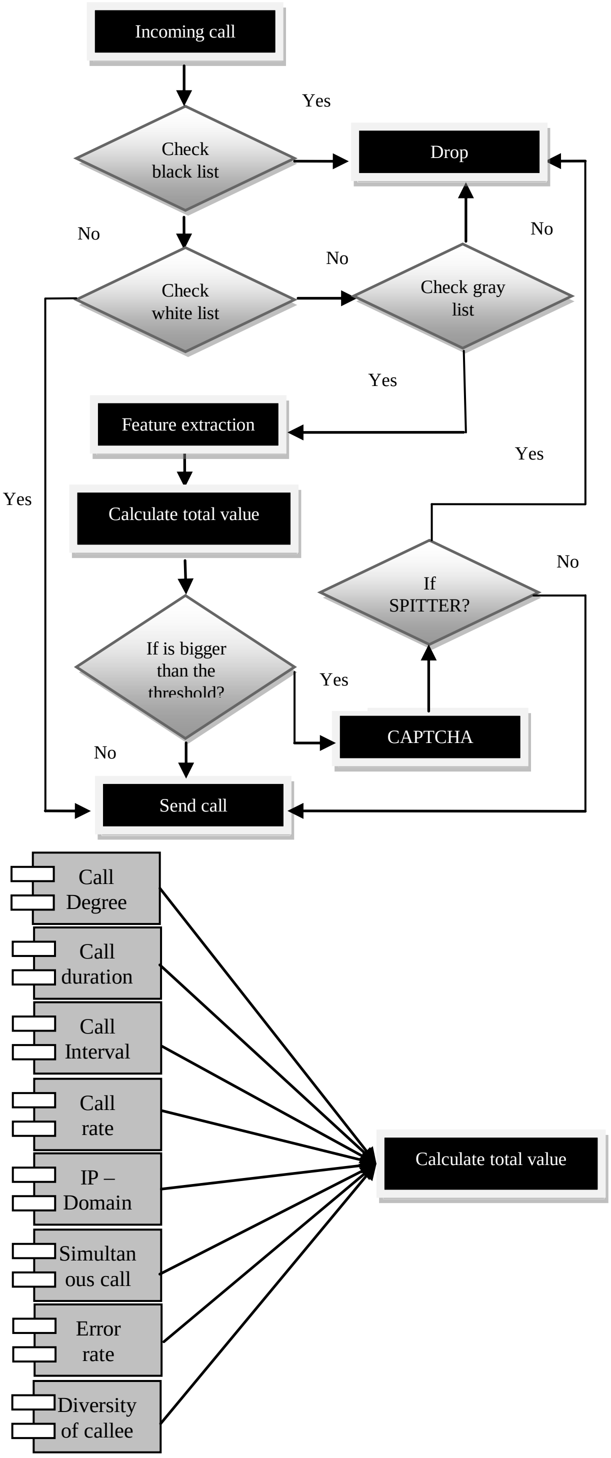

Fig. 1. Proposed framework.

- Simultaneous Calls: Occurrence of simultaneous calls to many users, precisely determines the suspicious behavior of a Caller.
- Diversity of Callers: A SPITTER attempts to cover a large amount of different Callees in a short time, while normal users generally call repetitive contacts and it is very rare that SPITTERS call repetitive contacts.
- Error Rate: SPITTERS encounter high volume of SIP errors including CANCEL packet and 404 errors. Errors are

generally occurred due to the interest of SPITTERs in propagating a message to several users (which may not be accessible or do not exist at all).

- Call degree: One of the features that are not used in SPIT prevention methods is investigation of input and output degree for users. SPITTERS usually make a lot of output calls while input calls of them are very low. Normal users have more double-faced calls.
- Other feature that is used in this proposed method is to check the Callee's IDs, IPs and domains which provide 8 combinations. For example if a user sends many calls with the same ID with different IPs, it could be suspicious of being a SPITTER.
- The last important feature in our proposed method is called Call interval. With checking features like the call period, duration, starting time etc., it could be realized that there is regularity in any of the Caller's calls. If there found any regularity then it means the Caller is sending automated calls.

B. *Proposed Framework*

Our framework considers weaknesses of other techniques and efficiently improves them. Using access lists plays an important role in SPIT detection since they are faster than other approaches and are easy to implement.

In our framework, we firstly check the incoming call within the lists. This will increase the detection speed by bypassing the feature extraction module which takes more time. We consider the decision made by the lists as deterministic, so updating it has to be performed with extra accuracy.

After checking the lists we extract the features of incoming call then calculate a total value. This value should be compared with a threshold. If it is greater than the threshold, the call is suspicious and needs more analysis hence we play the CAPTCHA for him in order to answer a question. If the answer is wrong, black list is updated and the call is dropped. On the other hand if number is below the threshold or he answers the CAPTCHA correctly, the call is sent to Callee. Fig. 1 shows the framework.

Total value for Caller is calculated according to the following formula:

$$\text{Total value} = \text{AVG (feature (1))} + \text{AVG (feature (2))} + \ldots + \text{AVG (feature (n))}$$

V. Simulation Results

We have simulated proposed mechanism on a pseudo-real database which is originally a combination of real call detail record from a VoIP company and six simulated SPITTERS with different complexity. They are deployed in MATLAB.

In Fig. 2 to Fig. 5 some of data samples are presented for both normal and SPIT calls.

Fig. 2 shows Call Rate for normal and SPIT users, it is obvious that SPITTERS have higher Call Rate than normal users since SPITTERS try to generate more calls in less time.

Call Duration of SPIT and normal users is shown in Fig. 3. When users answer a suspicious call, they immediately hang it up, since they find out that it is not a normal call, on the other hand we will have more duration for normal calls.

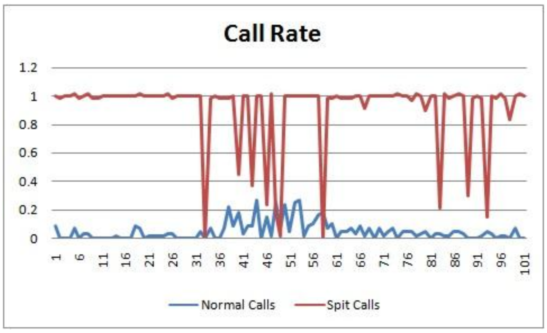

Fig. 2. Call rate.

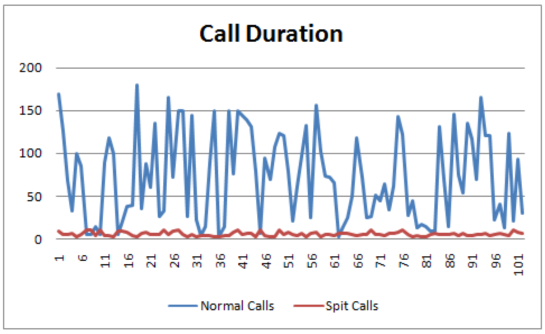

Fig. 3. Call duration.

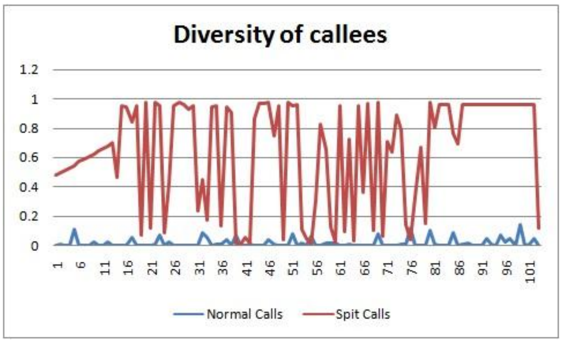

Fig. 4. Diversity of calles.

Diversity of Callees for normal and abnormal users is compared in Fig. 4, SPITTERS randomly create a list of users and they commonly tend to call a huge number of different users while normal users tend to call specific users hence SPITTERS have more diversity of Caller than normal users.

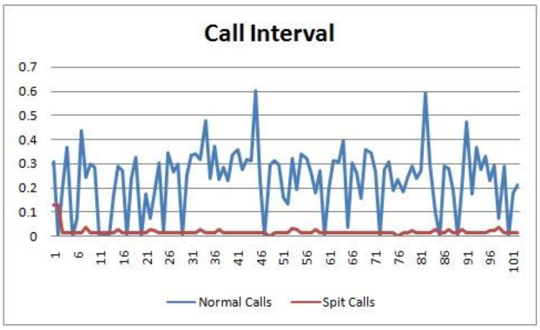

Fig. 5. Call interval.

Call interval is another feature for SPIT detection exactly inherited from [22]. The intertime distance for the calls generated by a SPITTER, usually follows a regular scheme (Especially for beginner SPITTERS). Hence, their calls have a low variance to mean ratio of previous intertimes.

In Fig. 5 we consider call interval for normal and SPIT users. SPITTERS usually call variety of users in a short time so they have very short regular distance between their calls while normal users show more irregular pattern between their calls.

Others features have binary values so they categorize in SPIT and non-SPIT categories. For example, Call degree is calculated according to the following formula:

$$\lim\left(\frac{\text{Input Degree}}{\text{Ouput Degree}}\right) = L$$

$$\begin{cases} \text{If } L \cong 0 \Rightarrow \text{SPIT} \\ \text{If } L \cong 1 \Rightarrow \text{non-SPIT} \end{cases}$$

As shown in formula above, if input degree to output degree ratio gets close to zero, it means number of output calls is more than input calls so we consider it as SPIT and if input degree to output degree ratio close to one, it means number of output calls is similar to input calls so we consider it as non-SPIT.

VI. Evaluation

Our aim is to decrease False Negative and increase True-Positive. The number of normal users that truly detected to be normal is called True-Positive, the number of normal users incorrectly detected to be SPIT is called False-Negative.

The final result is represented in Table III, which illustrates that each individual feature is unable to help the system in the detection progress. But a so called voting scheme provides an acceptable True-Positive and False-Negative rate.

TABLE III: True Positive and False Negative Rates

	Call Rate (%)	Call Duration (%)	Diversity of callees (%)	Call Interval (%)	Final value (%)
True-Positive	88.78	79.23	93.51	67.65	98.99
False-Negative	7.8	19.68	21.02	14.19	0. 64

VII. Conclusion

SPIT (Spam over Internet Telephony) is one of the major concerns in VoIP networks. Different detection solutions have been widely discussed in previous researches. Many of them categorized the incoming calls based on the extractable features in SIP headers but none of them has focused on the effective combination of these features.

In this paper, we extract important features that are effective in SPIT detection then we present a complete mechanism that includes all aspects. Results show that proposed mechanism is efficiently providing acceptable True-Positive and low False-Negative. Moreover, a comprehensive framework is proposed which applies lists and CAPTHCA.

References

[1] D. Sisalem *et al.*, *SIP security*, John Wiley & Sons, 2009.

[2] D. Gritzalis and Y. Mallios, "A sip-oriented spit management framework," *Computers & Security*, vol. 27, pp. 136-153, 2008.

[3] S. Dritsas *et al.*, "OntoSPIT: SPIT management through ontologies," *Computer Communications*, vol. 32, pp. 203-212, 2009.

[4] J. Quittek *et al.*, "Detecting SPIT calls by checking human communication patterns," in *Proc. IEEE International Conference on Communications*, 2007, pp. 1979-1984.

[5] S. Phithakkitnukoon and R. Dantu, "Defense against SPIT using community signals," in *Proc. IEEE International Conference on Intelligence and Security Informatics*, 2009, pp. 232-232.

[6] B. Mathieu *et al.*, "SDRS: A Voice-over-IP Spam Detection and Reaction System," *Security & Privacy, IEEE*, vol. 6, pp. 52-59, 2008.

[7] N. Chaisamran *et al.*, "Trust-based voip spam detection based on call duration and human relationships," in *Proc. IEEE/IPSJ 11th International Symposium on Applications and the Internet (SAINT)*, 2011, pp. 451-456.

[8] H. K. Bokharaei *et al.*, "You can SPIT, but you can't hide: Spammer identification in telephony networks," in *Proc. IEEE INFOCOM*, 2011, pp. 41-45.

[9] M. A. Azad and R. Morla, "Multistage spit detection in transit voip," in *Proc. 2011 19th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 2011, pp. 1-9.

[10] F. Wang *et al.*, "ADVS: a reputation-based model on filtering SPIT over P2P-VoIP networks," *The Journal of Supercomputing*, pp. 1-18, 2011.

[11] H. Hai *et al.*, "A SPIT Detection Method Using Voice Activity Analysis," in *Proc. International Conference on Multimedia Information Networking and Security*, 2009, pp. 370-373.

[12] H. Sengar, *et al.*, "Call Behavioral Analysis to Thwart SPIT Attacks on VoIP Networks," in *Proc. Security and Privacy in Communication Networks*, ed: Springer, 2012, pp. 501-510.

[13] A. Shahroudi *et al.*, "Full survey on SPIT and prediction of how VoIP providers compete in presence of SPITTERS using game-theory," in *Proc. 2011 IEEE International Conference on Computer Applications and Industrial Electronics*, 2011, pp. 402-406.

[14] S. Y. Park and S. G. Kang, "Labeling System for Countering SIP spam," in *Proc. 10th International Conference on Advanced Communication Technology*, 2008, pp. 1644-1646.

[15] M.-Y. Su and C.-H. Tsai, "A Prevention System for Spam over Internet Telephony," *Appl. Math*, vol. 6, pp. 579S-585S, 2012.

[16] Y. Bai *et al.*, "Adaptive voice spam control with user behavior analysis," in *Proc. 11th IEEE International Conference on High Performance Computing and Communications*, 2009, pp. 354-361.

[17] J. Seedorf, "Security challenges for peer-to-peer SIP," *Network, IEEE*, vol. 20, pp. 38-45, 2006.

[18] M. Stiemerling, *SIP: Protocol Overview*, ed: Radvision, 2001.

Mina. Aamanian was born in September 1988 in Mashhad, Iran. She received the B.S. degree in IT engineering from Sadjad University, Mashhad, Iran, in 2010. She is currently studying master of IT engineering at Imam Reza International University, Mashhad, Iran. She is the author of 1 international paper. Her research interests include Voice over IP technology, VoIP signaling & media security and SPIT detection in VoIP networks under supervision of Prof. Mohammad Hossein Yaghmaee Moghaddam. She works in IPPBX laboratory of Ferdowsi University as researcher and also she is working in Daysystem company.

Farideh Barghi was born in April 1988 in Mashhad, Iran. She received the B.S. degree in IT engineering from Azarbaijan Shahid Madani University, Tabriz, Iran, in 2010.

She is currently studying master of IT engineering at Imam Reza International University, Mashhad, Iran. Her research focuses on SPIT detection in VoIP networks under supervision of Prof. Mohammad Hossein Yaghmaee Moghaddam. She works in IPPBX laboratory of Ferdowsi University as researcher and also she is working in Khorasan Science and Technology Park.

Mohammad Hossein Yaghmaee Moghaddam was born on July 1971 in Mashhad, Iran. He received his B.S. degree in communication engineering from Sharif University of Technology, Tehran, Iran in 1993, and M.S. degree in communication engineering from Tehran Polytechnic (Amirkabir) University of Technology in 1995. He received his Ph.D. degree in Communication Engineering from Tehran Polytechnic (Amirkabir) University of Technology in 2000.

He has been a computer network engineer with several networking projects in Iran Telecommunication Research Center (ITRC) since 1992.

November 1998 to July 1999, he was with Network Technology Group (NTG), C&C Media Research Labs., NEC corporation, Tokyo, Japan, as visiting research scholar. Since 2000 he has been with the Computer Department of Ferdowsi University of Mashhad (FUM). He is the author of 3 networking books and more than 70 international papers. He is currently a visiting full professor at the Lane Department of Computer Science and Electrical Engineering, West Virginia University. His research interests are in Wireless Sensor Networks (WSNs), multimedia networking, traffic control, high-speed networks and fuzzy logic control.

Hossein Khosravi Roshkhari received his BSs and MSs from Ferdowsi University of Mashhad and is currently studying towards PhD at the same university. He was with GoldNet Engineering Group from 2009 and with Center for Research in Analog and VLSI microsystems dEsign (CRAVE) at Massey Univ., New Zealand in 2011 and 2012.

He is a technical manager at IP-PBX type approval laboratory at Ferdowsi University of Mashhad. His research interests include Voice over IP technology, VoIP signaling & media security and applications of AI in VoIP. He was lecturer at Unitec Univ., New Zealand and at IAUM Univ. at Mashhad, Iran.