SPIT Identification Criteria Implementation: Effectiveness and Lessons Learned

S. Dritsas, Y. Soupionis, M. Theoharidou, Y. Mallios, D. Gritzalis

Abstract While VoIP enables new means for communication, it may also provide a new way of transmitting bulk unsolicited messages and calls, namely SPam over Internet Telephony (SPIT). In this paper, we present the phases of a SPIT management process and we form a set of SPIT identification criteria, which are needed in order to characterize a call as SPIT, or a caller as spitter. Then, we present briefly the currently existing anti-SPIT frameworks, so as to examine which of the SPIT identification criteria is fulfilled by each framework, thus providing an insight on which criteria a technique should cope with, as well as how one can evaluate and combine existing approaches, in order to effectively mitigate SPIT. Finally, we implement a list of the criteria in our lab environment in order to examine the applicability of these controls in a Session Initiation Protocol (SIP) environment.

Stelios Dritsas
Dept. of Informatics, Athens University of Economics and Business, 76 Patission Ave., Athens, GR-10434, Greece, e-mail: sdritsas@aueb.gr

Yannis Soupionis
Dept. of Informatics, Athens University of Economics and Business, 76 Patission Ave., Athens, GR-10434, Greece, e-mail: jsoup@aueb.gr

Marianthi Theoharidou
Dept. of Informatics, Athens University of Economics and Business, 76 Patission Ave., Athens, GR-10434, Greece, e-mail: mtheohar@aueb.gr

Yannis Mallios
Information Networking Institute, Carnegie Mellon University, 4616 Henry St., Pittsburgh, PA 15213, USA, e-mail: imallios@andrew.cmu.edu

Dimitris Gritzalis
Dept. of Informatics, Athens University of Economics and Business, 76 Patission Ave., Athens, GR-10434, Greece, e-mail: dgrit@aueb.gr

Please use the following format when citing this chapter:

Dritsas, S., et al., 2008, in IFIP International Federation for Information Processing, Volume 278; *Proceedings of the IFIP TC 11 23rd International Information Security Conference*; Sushil Jajodia, Pierangela Samarati, Stelvio Cimato; (Boston: Springer), pp. 381–395.

1 Introduction

Voice-over-IP (VoIP) increasingly gains ground compared to traditional telephony. Its penetration and attractiveness is mainly due to its seamless integration with the existing IP networks, to its low-cost, and to the provision of sophisticated end-user services based on computer-based soft-phones. Currently, VoIP services drift towards the Session Initiation Protocol (SIP), due to its simplicity and its strong market acceptance. SIP is a protocol used for establishing communications between users, providing services such as voice telephony and instant messaging (IM) [14].

An identified threat to VoIP is the voice spam, referred to as Spam over Internet Telephony (SPIT). SPIT initiators, called spitters, use the IP network to generate bulk, unsolicited calls (or instant messages), mainly for commercial reasons. If SPIT prevalence becomes proportional to the one of spam, then the acceptance of VoIP will be encumbered. However, SPIT only recently received attention and only few solutions to it have been proposed (see section 4). Recent analyses show that SIP is more vulnerable to SPIT than it was initially estimated [5].

In this paper we argue that the effectiveness of any anti-SPIT technique is equally important to the actual technique itself. In this context we propose a set of SPIT identification criteria that will facilitate through their application a more concrete SPIT recognition and management process. Furthermore, we examine how the state-of-art antiSPIT mechanisms and frameworks handle the proposed criteria and finally, we present two different approaches that a VoIP system administrator could follow to implement these criteria in the domain that she is responsible for.

The paper is organized as follows: First, we illustrate some of the SIP features and present a macroscopic view of the SPIT management process. Then, we define a set of SPIT identification criteria needed to identify a SPIT call/message or a spitter. In Section 5, we briefly present existing anti-SPIT mechanisms. In section 6 we evaluate these mechanisms in terms of which SPIT identification criteria they cope with. Finally, we present two different ways of implement the predefined SPIT identification criteria and we conclude by providing the reader with some noteworthy remarks.

2 SPIT Phenomenon

SIP is an application layer protocol used to create, maintain, and terminate multimedia sessions. It supports five main services to multimedia communication: (a) user location, (b) user availability, (c) user capabilities, (d) session setup, and (e) session management. The basic SIP entities that support these services are User Agents (UA), which act as communication end-points, and SIP servers (proxies and registrars servers), which help and support the SIP sessions.

In this context, SPIT constitutes a new type of threat in VoIP environments. However, despite illustrating several similarities with email spam, there are certain differences between SPIT and spam, among them being the synchronous and real-time

nature of VoIP services, which hinder the adoption of email spam filtering techniques (i.e. Bayesian filters). Hence, new mechanisms should be adopted in order to handle effectively SPIT.

SPIT is defined as a set of bulk unsolicited phone calls or instant messages. Currently, three different types of VoIP spam forms have been recognized, namely: (a) *Call SPIT*, which is defined as bulk, unsolicited session initiation attempts to establish a multimedia session, (b) *Instant Message SPIT*, which is defined as bulk, unsolicited instant messages, known as SPIM, and (c) *Presence SPIT*, which is defined as bulk, unsolicited presence requests so as the malicious user to become a member of the address book of a user or potentially of multiples users.

The identified threats regarding SPIT are classified into four categories: (a) threats due to SIP protocol vulnerabilities, (b) threats due to the SIP protocol optional recommendations, (c) threats due to interoperability with other protocols, and (d) threats due to other (generic) security risks. These threats exploit specific SIP protocol vulnerabilities and can be used by a potential spitter in order to transmit SPIT calls and/or messages [5].

2.1 SPIT Management

The real-time nature of VoIP services led us to consider that it is more efficient to handle SPIT in the SIP signaling phase, than real-time filtering of a session (i.e. voice analysis). In general, a SPIT management process requires three distinct steps (see Fig. 1):

- ***Prevention***. This step prevents SPIT a priori, i.e., it impedes a potential SPIT message to be sent or a SPIT call to be established. In the context of SIP, prevention is responsible for blocking the spitter (caller) at her outgoing proxy. This requires a priori identification of SPIT, based on specific criteria. In order to be

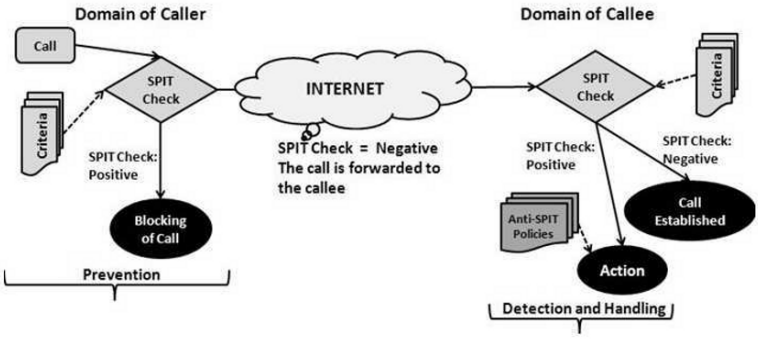

Fig. 1 A macroscopic view of SPIT Management

more efficient it should consider the overall SPIT policies that her domain has adopted.

- ***Detection***. This step detects a SPIT call or message when it reaches the callee's domain. It depends on pre-identified criteria and it is influenced by the preferences-interests of the callee, in terms of the attributes of the call or message, or the anti-SPIT policies of the domain of the callee.
- ***Reaction***. This step applies specific actions in case a call or a message has been identified as SPIT. These reactions, i.e. the application of specific anti-SPIT measures, are defined by the anti-SPIT policies adopted by the callee's domain.

3 SPIT Identification Criteria

SPIT management requires, first, appropriate criteria in order to identify SPIT calls and/or messages. In this section we present such a list of criteria, categorized according to their role in SPIT calls/messages. The same criteria can be used as detection rules, on both sides of a SIP session (i.e. outgoing or incoming proxies), according to their role in terms of handling SPIT. More specifically, when these criteria are used on an outgoing proxy (i.e. sender-caller's domain), we characterize them as preventive criteria, as they aim to prevent a call/message to leave the domain. When these criteria are applied by incoming proxies, they are characterized as detective criteria, as they aim to identify a SPIT call/message at the receiver point of a SIP session. The effectiveness of these criteria increases when they are applied in conjunction with strict identification and authentication mechanisms adopted by every SIP-participating domain. In this context, we propose two generic categories of SPIT identification criteria:

- ***SIP Message criteria***: This category includes the criteria that are related to attributes of SIP messages.
- ***SIP User Agent criteria***: This category includes the criteria that are related to attributes of a SIP User Agent.

Each one of the above generic categories is further analyzed into sub-categories, namely:

- Call and Message Origin as well as SIP participants' relationship (SIP User Agent oriented).
- Call and Message Patterns (SIP Message oriented);
- Headers' Semantics (SIP Message oriented).

A description of these criteria is presented in the sequel.

3.1 Call and Message Origin and SIP participants' relationship Criteria (SIP User Agent Oriented)

This category includes criteria that examine the characteristics of a SIP session, regarding the SIP addresses of the sender/caller (i.e. SIP URI or IP address), as well as the domain the session was initiated in[1]. Furthermore, through this analysis the relationship of the participants of a SIP session is examined, i.e. whether the caller/sender is trusted by the callee/receiver. Typical examples include whether a caller is known to the callee (included in his address book), whether she is included in a white list or contrary she is blacklisted.

- *Caller SIP URI*: It detects and analyzes the SIP URI of the sender of a call/message, so as to determine if she is a potential spitter or not.
- *Caller IP Address*: It analyzes the IP address of the sender/caller so as to characterize her as spitter.
- *Caller Domain*: It analyzes the identity of the domain of the caller (sender), which is determined either by SIP URI of the caller, or through DNS lookup from her IP address. If the identity of the domain is a well-known SPIT source, then the call or the message is characterized as potentially SPIT.

3.2 Call-Messages Patterns (SIP Message Oriented)

This category includes criteria that analyze specific call or message characteristics or patterns, in order to determine whether a call (message) is a possible SPIT.

- *Path traversal*: A call or a message might pass through many intermediates before reaching its final destination. This path is denoted in the Via header. Thus, if in the Via header a SPIT domain is recognized, the call or the message may be a potential SPIT.
- *Number of calls-messages sent in a specific time frame*: It analyzes the number of calls (messages) made in a specific time period by a user. If this number is above a specific pre-defined threshold then the call (message) is characterized as a possible SPIT call.
- *Static calls' duration*: If the calls initiated by a single user have a static duration, then the user is a potential spitter and it is possible to use an automated script (e.g. bot software) in order to make the calls.
- *Receivers' address patterns*: If the receivers' addresses follow a specific pattern (e.g. alphabetical SIP URI addresses), then the call (message) is flagged as SPIT.

[1] The specific controls require a database that stores the source of well-known spitters' domain or specific spitters identities (SIP URIs).

- *Small percentage of answered/dialed calls*: It indicates the number of successful call completions from this caller per a pre-defined time period, which is relative to the number of failed ones.
- *Large number of errors*: When a user send a large number of INVITEs and the SIP protocol returns a large number of error messages (e.g. 404 Not Found) then it is probable this user be a potential spitter, therefore the calls made by her/him are blocked.
- *Size of SIP Messages*: In this case a set of SIP messages sent by a user to other users is analyzed. If those messages have a specific size then it is very possible to be sent by a "bot" software, therefore the call is characterized as SPIT.

3.3 SIP Headers' Semantics (SIP Message Oriented)

This category includes criteria that identify a SPIT call or message through a semantic analysis of the contents of the SIP messages. Through the analysis one can apply well-known anti-spam techniques (e.g. Bayesian filters), in order to determine if a call/message is SPIT.

These particular criteria are further categorized, according to the different parts of SIP messages that could be used. These are: (a) a message's headers, (b) a message's body, and (c) the reason phrases of a message.

In addition, we have identified three possible types of SPIT that could be injected in a SIP message, namely: (a) text SPIT injected in a header field, (b) media SPIT carried in the message's body, and (c) hyperlink to a SPIT message injected in a header field).

Tables 1 to 3 depict the specific SIP header fields that can be used for a detailed semantic analysis, so as to detect a SPIT call or message alongside with the type of the SPIT that could be sent. Hence, Table 1 presents the header fields of the SIP request or response messages that should be examined in order to check if they include SPIT content. For instance, the header *Subject* might contain a suspicious text, i.e. the word pornography, which in most cases might be considered as SPIT. Moreover, the *Alert-Info* header might include a hyperlink that directs a user to a specific site used for promotional-commercial reasons.

Table 2 presents the types of SIP messages that could contain a body field. This field should be examined as it may include suspicious content, characterized as SPIT. The message types are grouped in Request and Response Messages. The SPIT conained in the message body can be text, media or hyperlink.

Table 3 presents the Reason Phrases of Response Messages that could be used by a malicious user so as to generate SPIT message. More specifically, the Reason Phrases may consist of plain text or hyperlink, which forms the SPIT message sent to the receivers.

Table 1 SIP Headers that could include SPIT content

Header Fields	SPIT Type	Request Messages	Response Messages
Subject	Text	✓	✓
From	Text	✓	✓
Call-Info	Hyperlink	✓	✓
Contact	Text	✓	✓
To	Text	✓	✓
Retry After	Text	✓	✓
Alert-Info	Hyperlink	✓	✓
Reply To	Text	✓	–
Error-Info	Hyperlink	–	✓
Warning	Text	–	✓
Header Fields related to SIP messages' bodies *notcarryingSPIT*"*directly*"			
Content-Disposition	Displayed Message Body	✓	✓
Content-Type	Displayed Message Body	✓	✓

Table 2 Request-Response Messages that could include SPIT content

Message Type	**Message**
Request Messages	INVITE ACK
Response Messages	180 Ringing 183 Session Progress 200 OK 300 Multiple Choices 380 Alternative Service 488 Not Acceptable Here 606 Not Acceptable

Table 3 Request-Response Messages that could include SPIT content

Response Messages Possibly Carrying Reason Phrases
182 Queued
183 Session Progress
200 OK
400 Bad Request
480 Temporarily Unavailable
484 Address Incomplete

4 Anti-SPIT Mechanisms Overview

As mentioned, SPIT may influence the future use and adoption of the VoIP technology. So far, some general frameworks from the email spam paradigm have been discussed as candidates for SPIT handling [13]. Furthermore, some of them appear to be basic building blocks of the anti-SPIT architectures that have been proposed in the literature. In the sequel, we discuss the anti-SPIT architectures that have been proposed so far.

AVA (Anonymous Verifying Authorities). The Anonymous Verifying Authorities approach, presented in [2], is based on the introduction of a "call-me-back" scheme and the use of two new entities, namely: (a) the Mediator and (b) the Anonymous Verifying Authority (AVA). The authors try to mitigate SPIT by anonymously blocking unwanted calls through AVA and the Mediator. Thus, in the case of not call establishment, the caller is not aware for the existence of the callee.

Anti-SPIT Entity. A network-level entity, placed in the edge of the network, is proposed in [8]. The role of this entity is to filter and analyze the transmitted SIP packets, and to detect SPIT according to certain criteria. By using these criteria, a weighed sum is introduced, namely spitLevel, which serves as a threshold. If the spitLevel is exceeded specific actions are performed depending on the policies adopted by the callee's domain. Experimental data are provided.

Reputation/Charging Mechanism. The work in [13] proposes two techniques for handling SPIT. The first is based on reputation builds trust within different SIP communities and uses the resulting trust networks for detecting SIP spam. The second is a variant of the payment at risk proposal. Implementation details are not provided by the authors.

DAPES (Domain-based Authentication and Policy-Enforced for SIP). In this framework, any SIP-based communication passes through two stages of verification; namely, verification of the caller's identity, and mutual authentication of the participated proxies alongside with verification of the outbound proxy [17].

PGM (Progressive Multi Gray-Levelling). The approach proposed in [4], stems from the antiSPAM framework graylisting. Accordingly, it calculates and assigns a non permanent gray level for each caller, in order to check if a message is SPIT or not. This level is calculated based on previous call patterns of a particular caller. Depending on the level's value, appropriate actions are taken.

Biometrics Approach. In [1], the authors propose the use of global servers that bind users' identities to personal data; they select biometric data, such as a person's voice. The proposal is based on the concept of binding identities to persons that cannot change globally. User interference and threats taken into account are also mentioned.

RFC 4474. An end-user authentication scheme is discussed in RFC 4474 [11], based on Public Key Infrastructures (PKI) and Certificate Authorities (CA). Although this approach is not oriented specifically towards SPIT handling, the identity control mechanism is useful for controlling SPIT. Two new SIP header fields are used and their manipulation is done only by proxy servers within the domain of the calling UA, through appropriate authentication and certificates.

SIP-SAML. The approach presented in [18] uses the Security Assertion Markup Language (SAML) for SIP authentication and authorization through asserted traits. The authors aim at a strict identity control accomplishment, in order to prevent spitters from changing their identity frequently.

DSIP (Differentiated SIP). In [6], an extension to SIP is proposed. It tries to handle SPIT through the classification of callers into three categories of lists, namely: white, who are legitimate callers, black, who are spitters, and grey list, who are not yet classified. Through this classification of users, the handling of calls is conducted accordingly. When the caller is unknown, a human verification test is imposed, in order to prove that she is not a SPIT automated machine.

VoIP Seal. The work in [9] presents a system that operates in two stages. During the first stage, modules that are not transparent to the caller, examine the call. Each module of the first stage contributes a score in [-1,1], where high score corresponds to a high probability that the call is SPIT. Each module is associated with a weight, and the final score is compared with two thresholds. If the score is within acceptable threshold range, then the call passes to the second stage of checking the call. This stage includes modules that require interaction with the user. For instance, they could be a Turing test that checks whether the caller is spitter or not. If this test fails, the call is rejected.

VSD (Voice Spam Detector). The [3] framework combines many of the anti-SPIT approaches presented in [15]. The system is a multi-stage SPIT filter based on trust and reputation, and uses feedback between the different stages. Presence filtering, the first step, depends on the current status of the callee. The next step, that is the traffic pattern filter, analyzes the rate of incoming calls. This step is followed by the white/black lists' filtering. Bayesian filtering is the fourth step, where a call is checked regarding the behavior of the participated entities. Finally, reputation techniques are used to check the acceptance of a call based on social relationships that an end- user maintains with other entities in the VoIP environment.

5 Compliance of SPIT mechanisms to Identification Criteria

In this section we identify the SPIT identification criteria which have been used by the aforementioned mechanisms. Our analysis, in conjunction with the analysis presented in [7] provides the reader with a point of reference, in terms of which mechanisms should be selected in a specific context. In this context, Figure 2 presents which of the mechanisms takes into account the SPIT identification criteria we defined. For this purpose we took under consideration only an abstract description of each mechanism as implementation details are not fully discussed and described, in their relative publications. Furthermore, we do not consider whether the mechanisms meet the criteria well or not, but we rather provide the mere existence of each criterion in the mechanisms' description. For example, in the description of Reputation/Charging mechanism, the use of Black and White lists requires the existence of a way to identify and handle users, either by SIP URI, IP address or even domain

of origin. However, as something like that is not explicitly mentioned, we put the appropriate negative value in the table.

Furthermore, the table can be used as a reference to choose the appropriate mechanism for SPIT handling in a given context. For example the call and message patterns might be costly to implement, in terms of data gathering and analysis, thus mechanisms that focus on and fulfill the other criteria might be of preference.

Finally, the table can be read as a concentrated area of further research directions regarding anti-SPIT countermeasures. Some of the questions that one can answer using the table include how can a particular mechanism contribute in terms of prevention, detection or handling of SPIT, which combinations of techniques should someone use in order to fight SPIT more effectively, etc.

6 Implementation

A key question regarding the proposed criteria is whether they can be applied on a SIP environment. In order to examine their applicability, we first implemented the following test computing environment, which is depicted in Fig. 3. It consists of a SIP Proxy Server, which is established in our laboratory environment. The SIP server application is a scalable and reliable, open source software called SIP Express Router (SER 2.0) [16]. It can act as a SIP registrar, proxy, or redirect server. We have extended its functionality to support our implementation of the above mentioned criteria. All the laptops and the PCs are equipped with soft-IP-phones (X-lite), which can use the SIP server in order to establish a call.

Having the above testbed in a full functional status, we implemented the proposed identification criteria so as to examine their applicability in real VoIP settings. From

SPIT Identification Criteria / Anti-SPIT Mechanisms	Call-Message Origin Caller/Callee Relationship				Calls-Messages patterns							SIP Headers Semantics
	Caller SIP URI	Caller IP address	Caller domain	White/Black list	Path of message	Number of Calls	Calls duration	Sequential Call Numbers	Dialed/answered calls	Large number of errors	Size of SIP message	Message Headers, Bodies, and Reason Phrases
AVA	☑	–	–	–	–	–	–	–	–	–	–	–
Anti-Spit Entity	☑	☑	–	☑	–	☑	☑	☑	–	☑	–	–
Reputation/Charging	–	–	–	☑	–	–	–	–	–	–	–	–
DAPES	☑	☑	☑	–	☑	–	–	–	–	–	–	–
PMG	☑	–	–	☑	–	☑	–	–	–	–	–	–
Biometrics	–	–	–	–	–	–	–	–	–	–	–	–
RFC 4474	☑	–	☑	–	–	–	–	–	–	–	–	–
SIP SAML	☑	–	☑	–	–	–	–	–	–	–	–	–
DSIP	☑	–	☑	☑	–	–	–	–	–	–	–	–
VoIP Seal	–	–	–	☑	–	–	–	–	–	–	–	–
VSD	☑	☑	☑	☑	☑	☑	–	–	–	–	–	–

Fig. 2 A macroscopic view of SPIT Management

the list of the criteria, presented in section 3, we selected the (a) Call and Message Origin and (b) SIP participants' relationship criteria, as well as the (c) SIP Headers' Semantics criteria. Regarding the Call-Messages Patterns category, we implemented only the path traversal (path of message), because the remaining ones require historical and statistical data in order to generate metrics and define thresholds. For instance, the numbers of calls criterion requires the historical logs per caller which might introduce modifications in the setup of our environment.

We have used two different approaches to put in practice the criteria. In the first technique, we alter the main configuration file of the SIP Server. In the second one, the main parameters of the criteria are stored into an external MySQL database (ver. 5.0) and for each SIP message we query the database in order to find out if it is SPIT message or not. MySQL database is also used by SER for storing users, as a part of the typical setup of the SER server. In the following, we present two implementation examples and then we compare the two techniques.

6.1 Implementation with configuration file

The SER configuration file consists of the main SIP Server attributes and the routing rules of the SIP messages. The SPIT criteria are applied by adding a small portion of additional code in the configuration file for each criterion, which is identified by the SIP server administrator.

An example of an implemented criterion, which is mentioned in paragraph 3.3, is stated below. *Example code 1* shows a SIP message, in which the Error-Info (highlighted) contains a Hyperlink. Therefore Bob's SIP proxy server will reject the incoming call.

The proposed addition in the configuration file so as to discover this vulnerability is described in *example code 2*. The first line is used to discover the proposed criteria, the second line is used to write in a log file the reason for which the message was

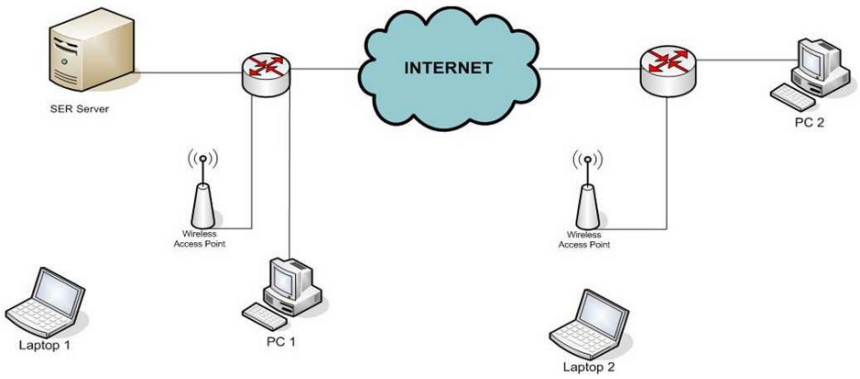

Fig. 3 A macroscopic view of SPIT Management

rejected, the third line is used to send a response SIP message to the caller explaining the cause of the rejection and the forth line is used to terminate the connection.

***Example Code 1:** SIP Message (Error-Info header field contains a hyperlink)*

```
1: SIP/2.0 200 OK
2: Via: SIP/2.0/UDP pc1.example.com
3: To: "Bob" <sip:bob@example.com>;tag=987
4: From: "Alice" <sip:alice@aueb.gr >;tag=123
5: Call-ID: 6543219999@172.1.2.2
6: CSeq: 1 INVITE
7: Contact: <sip:bob@example.com>
8: Content-Type: application/sdp
9: Content-Length: 200
::
21: Error-Info: <http://www.sell.com/yourshoe.jpg>
```

Example Code 2: Error-Info criterion script (part of SER conf.file)

```
1: if (search("^Error-Info:\s<http://.*"))
2: {
3: log("LOG: alert: someone trying to send an
      http link through Error-Info\n");
4: sl_send_reply("476", "No Hyperlink Text is
      permitted through Error-Info" );
5: break;
6: \};
```

6.2 Implementation with MySQL Database

The MySQL database is used to store all the parameters which assist to identify a possible SPIT. For example it stores all the domains and URIs of callers for which it is decided whether they are spitters or not. Therefore, each newly received SIP message is partially passed to an external script which performs a query to the database and checks if the message violates any of the given SPIT rules.

A script, which finds out if the user's URI (mentioned in paragraph 3.1.) is acceptable, appears in the sequel (*example code 3*).

Example Code 3: External script accessing database

```
#!/bin/sh
m=`echo $1 | sed -e 's/^sip://'`
num=("echo 'SELECT count(*) FROM users
     WHERE user_uri=\"$m\";'
| mysql -u ser -h localhost --password=heslo -D ser")
if [ $num != 0]; then
    exit 0
else
    exit 1
fi
```

6.3 Implementation Results and Comparison

The main advantage of the first technique is the speed of (a) handling the SIP messages and (b) deciding whether the message is SPIT or not. This occurs because, for routing every SIP message, the configuration file is accessed. On the other hand, it is really complex to insert a new SPIT criterion. For example, if the administrator decides to reject all incoming calls from a certain domain, he has to find out the exact position in the configuration file to place the script and afterwards he has to restart the SER server in order this modification to take effect.

The second method helps the administrator to add and modify values of SPIT criteria without the reloading of the SER instance being mandatory. The main drawback of this method is the time overhead as it has to access the database for every message and actually execute a query for each criterion in each message.

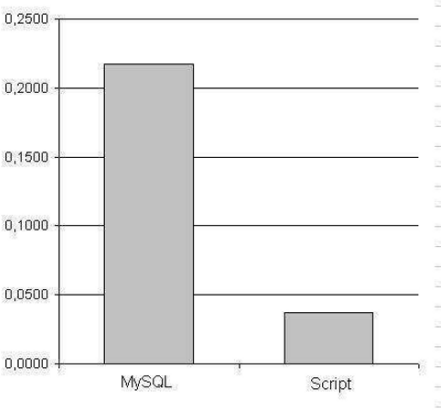

Fig. 4 Performance Comparison of two suggested methods

The performance comparison of the two techniques is presented in the figure 4, where we examined the time required for extracting and checking the SIP URI address of a SIP packet. The related time was 0,21748 sec for database script and 0.0371 for SER configuration script respectively.

7 Discussion and comments

VoIP technology and SIP raised significant concerns, as to whether SPIT phenolmenon will be equivalent to the current spam prevalence. In order to address and evaluate these concerns, we provided a macroscopic view of SPIT management, alongside with an extensive list of SPIT identification criteria that can be used by anti-SPIT mechanisms in the prevention and detection stages of SPIT management.

VoIP infrastructures have recently gained a (still) small, but recognizable market share. Thus, only recently, and prior to the SPIT phenomenon prevailing, some anti-SPIT mechanisms have been suggested. The majority of them focus on the prevention, detection, and handling stages of SPIT management. Most of them seem not to take into account the results of an appropriate threat and vulnerability analysis regarding SPIT, thus SIP protocol vulnerabilities are usually not considered.

On the other hand, the proposed anti-SPIT mechanisms aim at fulfilling qualitative and quantitative criteria. In this paper we used a two-fold evaluation framework. First, we defined a set of parameters that each mechanism should address in order to counter SPIT efficiently, and we identified how each class should be evaluated, in terms of effectiveness. Second, we analyzed which of the SPIT identification criteria each SPIT mechanism takes into account. Finally, we implement two methods of discovering the possible SPIT messages. It is clearly demonstrated that not only it is achievable to put in practice the proposed criteria, but also that the methods are considerably effective since, as they check every SIP message for possible SPIT attributes. A possible extension of the proposed implementation would aim at taking into account the criteria presented in section 3.2.

Finally, the proposed evaluation framework provides insight on how the effectiveness of a mechanism can be evaluated, as well as how combinations of relevant mechanisms should be selected, in order to effectively mitigate SPIT in a given context. In this context, we are planning to implement an automatic solution that allows us to evaluate each anti-SPIT mechanism based on the criteria and choose the best one.

Acknowledgements This work has been partially performed within the SPIDER (COOP-32720) project, which is partly funded by the European Commission under Framework Programme 6. We would like to thank our partners in the consortium for their constructive co-operation.

References

1. Baumann R., Cavin S., Schmid S.: Voice Over IP - Security and Spit. Swiss Army, FU Br 41, KryptDet Report, Univ. of Berne (2006)
2. Croft, N., Olivier, M.: A Model for Spam Prevention in Voice over IP Networks using Anonymous Verifying Authorities. In: Venter, H.S. et al. (eds.) Proc. of the 5th Annual Information Security South Africa Conf., South Africa (2005)
3. Dantu R., Kolan P.: Detecting Spam in VoIP Networks. In: Proc. of Steps to Reducing Unwanted Traffic on the Internet Workshop, USA (2005)
4. Dongwook S., Jinyoung A., Choon S.: Progressive Multi Gray-Leveling: A Voice Spam Protection Algorithm. IEEE Network, 20(5), 18-24 (2006)
5. Dritsas S., Mallios J., Theoharidou M., Marias G., Gritzalis D.: Threat Analysis of the Session Initiation Protocol Regarding Spam. In Proc. of the 3rd IEEE Int. Workshop on Information Assurance (26th IEEE International Performance Computing and Communications Conference), IEEE Press, pp. 426-433, New Orleans (2007)
6. Madhosingh B.: The Design of a Differentiated SIP to Control VoIP Spam. Technical Report, Computer Science Department, Florida State University (2006)
7. Marias G., Dritsas S., Theoharidou M., Mallios J., Gritzalis D.: SIP vulnerabilities and anti-SPIT mechanisms assessment. In Proc. of the 16th IEEE Int. Conf. on Computer Communications and Networks (ICCCN '07), IEEE Press, Hawaii, pp. 597-604 (2007)
8. Mathieu, B. et al.: Spit Mitigation by a Network-Level Anti-Spit Entity. In: Proc. of the 3rd Annual VoIP Security Workshop, Germany (2006)
9. Niccolini S.: Spit prevention: State of the art and research challenges. Network Laboratories, NEC Europe, Germany (2006)
10. Park S., Kim J., Kang S.: Analysis of applicability of traditional spam regulations to VoIP spam. In: Proc. 8th Int. Conf. of the Advanced Communication Technology, Phoenix Park, Korea (2006)
11. Peterson J., Jennings C.: Enhancements for Authenticated Identity Management in the Session Initiation Protocol. RFC 4474 (2006)
12. Rebahi Y., Sisalem D.: SIP service providers and the spam problem. In: Proc. of the Voice over IP Security Workshop, Washington, USA (2005)
13. Rebahi, Y., Sisalem, D., Magedanz T.: SIP Spam Detection. In: Proc. of the Int. Conf. on Digital Telecommunications, pp. 29-31, France (2006)
14. Rosenberg, J. et al.: Session Initiation Protocol. RFC 3261 (2002)
15. Rosenberg J., Jennings C.: The Session Initiation Protocol (SIP) and Spam. draft-ietf-sipping-spam-03 (2006)
16. SER server version 2.0, Available via www.iptel.org/ser. Cited 10 Jan 2008
17. Srivastava K., Schulzrinne H.: Preventing Spam For SIP-based Instant Messages and Sessions. Technical Report, University of Columbia (2004)
18. Tschofenig H. et al.: Using SAML to Protect the Session Initiation Protocol. IEEE Network, 20(5), 14-17 (2006)