SIP Service Providers and The Spam Problem

Y. Rebahi, D. Sisalem
Fraunhofer Institut Fokus
Kaiserin-Augusta-Allee 31
10589 Berlin, Germany
{rebahi, sisalem}@fokus.fraunhofer.de

Abstract

The Session Initiation Protocol (SIP) is used for managing multimedia sessions in the Internet. As an emerging standard gaining more and more acceptance within the IT community, SIP will probably be the target of spammers. We propose, in this paper, a reputation-based mechanism that builds trust between users within a SIP community and prevents spammers to carry out attacks. Our technique uses a social networks approach enhanced with some reputation ratings. A metric for computing reputation is defined and an algorithm allowing the users to evaluate each other is also presented.

1 Introduction

Spam refers in general to any unsolicited communication. The latter may hide actions that vary from marketing advertisements to some computer virus spreading. Spammers intend to create an environment of "untrust" in the provided services as well as the tools used for the provision. For instance, some statistics [8] reveal that 52% of email users say spam has made them less trusting email and 25% say that the volume of spam has reduced their usage of email. Spam appears as a real threat that the service providers try to combat using in particular some filtering mechanisms. These filters are not always efficient as they have to make sure before discarding any message that the latter is a spam message with an extremely high confidence. In addition, these filters also need to be maintained and updated regularly.

Reputation systems ([15],[16]) are mechanisms that have been used in many different areas to build trust between members of a certain community. As an example, one can mention eBay which is an electronic community where the members can be rated according to the performed commercial transactions. The reputation systems are in general lightweight and can be maintained easily.

In this paper, we describe a mechanism that uses the reputation concept for building trust between users in a SIP community. Our mechanism allows the users to rate each other according to some predefined criteria. Based on this rating, a "SIP social network" is developed. This allows to identify the SIP spammers and discard them from the network. Our scheme could also be used in concert with the existing filtering techniques to prevent spam problems within SIP communities.

The paper is organized as follows: Section 2 presents an overview of the SIP spam problem and the existing solutions to deal with. Section 3 describes the proposed architecture as well as its different functionalities. In section 4, some issues that will be addressed in the future are listed and finally section 5 concludes the paper.

2 Spam problem in SIP

2.1 SIP Overview

The Session Initiation Protocol (SIP) is an application-layer protocol defined by IETF for managing multimedia sessions in the Internet [1]. A session can be established between two end-users or more and can involve IP phone calls, conferencing and messaging. SIP is handled through messages such as INVITE for initiating a session and BYE to terminate it. The main components in a SIP network are briefly described below ([1], [2]),

User Agents (UAs): these are the end devices in a SIP network. They can originate SIP requests to establish a media session and send and receive media. A user agent can be either a SIP phone or a SIP client software running on a PC.

Servers: these are intermediate SIP entities in a SIP network that assist the user agents in establishing media sessions and some other functions. SIP servers are three categories: proxies, redirect servers and registrars.

Location servers: a location server is a database where users information such URLs, IP addresses, scripts and other preferences are stored. A location server may also contain routing information such as locations of proxies, gateways and other location servers.

The SIP standard is more and more accepted by the IT community and currently is ubiquitous in the market through a large row of services. For instance, the new Messenger for Windows launched by Microsoft, includes SIP-based telephony, presence, instant messaging and voice/video communication. SIP is increasing its popularity in the mobile world too: in the 3GPP consortium for 3G mobile networks, SIP is the chosen signaling protocol and Nokia has announced SIP support for its Series 60 platform of mobile terminals. The SIP protocol has shown a strong acceptance by the market as some service providers offer already free and non-free SIP-based VoIP and Instant Messaging services. Among these SIP providers, one can mention earthlink [3] and iptel.org [4].

To provide satisfactory services, SIP providers need to deal with unsolicited or spam communications. The spam problem in email and instant messaging (IM) has conducted the email or the IM users to trust less these tools and consequently reduce their usage. SIP has not been yet the target of this kind of attacks, but it seems it is only a matter of time [5].

2.2 SIP Spam forms

As a new emerging standard, SIP might be also the target of some Spam attacks. As a consequence, identifying SIP spam and the mechanisms to deal with, is a very crucial task before the problem arises. SIP spam can take one of the following forms, for more details, we refer to [5],

- *call spam :* this is the case of unsolicited messages for establishing voice, video or IM session. The spammer proceeds to relay his message over the real time media. This form is the usual way used by telemarketers
- *IM spam:* this form is similar to email spam, unsolicited IMs whose content contains the message that the spammer is seeking to convey. The SIP MESSAGE request will be used here but also some other messages such as INVITE with text or html bodies

- *Presence spam:* this spam is similar to the previous one, i.e unsolicited presence (subscribe) requests are sent to get on the buddy list of a user and send him IM or initiate other forms of communications

In [5], a bunch of solutions to SIP spam is also suggested. In general, these solutions are some mechanisms developed to deal with email spam and which could be adapted to SIP spam because of some similarities as mentioned earlier. Among these solutions, one can mention, content filtering, white lists, Consent-based communications and identity authentication [5]. Unfortunately, if these solutions can be applied to the SIP spam problem, they seem to be insufficient and need to be combined with each other or with some other techniques to provide robustness.

3 The SIP Social Networks Approach

3.1 Graphs Structure Overview

A graph is a finite set of nodes (or vertices) connected by links called edges (or arcs). A directed graph is a graph where each edge is an ordered pair of nodes (u,v) representing a directed connection from u to v. The *out-degree* of a node u is the number of the distinct edges $(u,v_1)...(u,v_k)$ and the *in-degree* is the number of the distinct edges $(v_1,u)...(v_k,u)$. The degree of a node u is the sum of the in-degree and the out-degree values. A path from node u to node v is a sequence of edges (u,u_1), (u_1,u_2), ... (u_k,v). Note that a path from u to v does not imply a path from v to u.

3.2 SIP Social Networks

In general, a network refers to a set of items called vertices (or nodes) with possible connections between them, called edges. A social network is a set of people with some pattern of contacts or interactions between them. These patterns can be friendship relations between individuals, business relationships between companies, etc. Systems taking networks or social networks forms are called graphs in the mathematical literature (see section 3.1).

Let's consider now a SIP Provider that offers SIP based services such as VoIP calls and Instant messaging. Let's assume that any user of these services is managing a contact list to keep track of his "friends" that they are allowed to communicate with him. Usually, buddy lists are used for the Instant Messaging service and phone books are used for the telephony services. However, along this paper, contact lists will be used as a general terminology to describe lists of a user's "friends" that are allowed to communicate with him. Let's also assume that any SIP user is able to assign a reputation value to all the persons figuring on his contact list. These reputation values simply refer to how trustful the assigned persons are. It turns out when we combine all these scored contact lists, a large reputation network is generated. In fact, this network is constructed by considering first the owner of the contact list and the users figuring on this list as nodes connected with some scored edges. These sub-networks are connected to each other when there is a possibility to connect them. Figure 1 depicts how two users A and B get connected to each other through their common "friends".

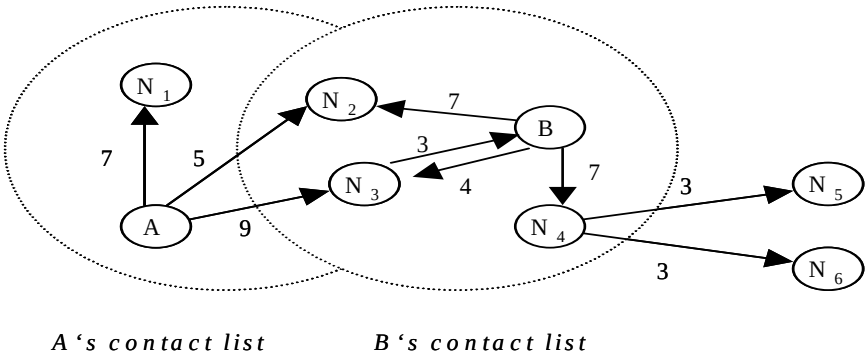

Figure 1: SIP Reputation Network

In the sequel, we will use graphs or "SIP social networks" when talking about the networks obtained by merging all the contact lists (in the way described above) of the SIP users. In doing so, we will avoid any confusion with the SIP networks (infrastructure) terminology described in section 2.1.

3.3 Architecture

First of all, the objective of this architecture is to provide a mechanism based on reputation to help detecting unsolicited calls or messages within the users community of a given SIP Service Provider. In fact, it may happen that a user from this community misuse one or more of the provided services by sending unsolicited messages or establishing VoIP calls. The intentions behind this misuse can be business purposes or malicious actions. For detecting SIP spam, our system does not require anything on the user's side except that when the users create their contact lists, they must score the persons on these lists according to an evaluation criteria that the Service Provider has predefined.

Figure 2 depicts the different components that our system comprises. These components as well as their functionalities will be discussed below.

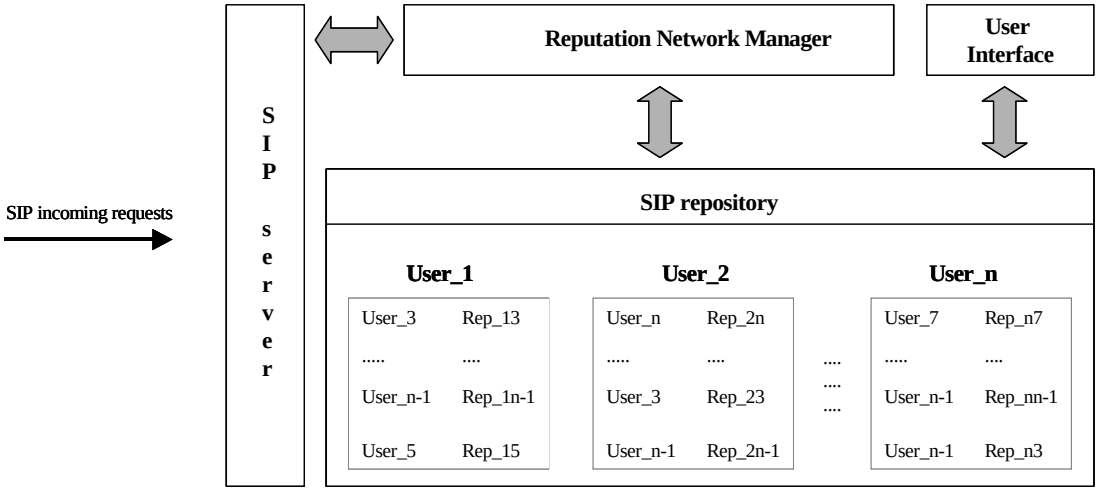

Figure 2: Network architecture

3.3.1 SIP repository

The SIP repository is a database where the contact lists of the different users are stored. The contact list records are of the form described in Figure 2. The first column of these records comprises the users identifiers, which could be the names of these users or their SIP URLs. However, the second column includes the reputation values that the owner of the list has assigned to each member of his contact list. The contact list records are created and updated by the user through a web interface or another suitable tool.

3.3.2 SIP server

As mentioned in section 2.1, a SIP server is an intermediate entity that assists the SIP user agents in establishing the multimedia sessions. In our architecture, the SIP server is the main component as the multimedia services such as VoIP calls and Instant Messages are handled through it. In fact, a spammer needs to issue a SIP request towards the users that the spammer wishes to be in contact with. The SIP request can be for instance the SIP INVITE or the SIP MESSAGE. The SIP server intercepts the mentioned request and before performing the corresponding action, it forwards the request sender identifier and the request destination identifier information to the Reputation Network Manager (RNM) in order to check whether this request is an unsolicited message. Once the SIP server receives the answer from the RNM, it behaves accordingly to treat the received request.

3.3.3 Reputation Network Manager

The first functionality that the RNM has to perform is building the SIP social network. The RNM has the access to the SIP repository, so, based on the stored contact lists records, it can construct the reputation network in the way described in section 3.2. As the contact lists might be updated, the RNM has to access regularly the SIP repository and update consequently the SIP social network.

The other functionality that the RNM has to carry out is the reputation computing. This simply means that upon receiving the SIP request information from the SIP server, the RNM computes a reputation value related to this request and compare it to a predefined threshold that will be discussed in section 3.3.3.2. If this reputation value is greater than the threshold, the RNM will inform the SIP proxy to process the request, otherwise, the SIP server will reject it.

3.3.3.1 Reputation computing algorithm

First of all, the RNM needs to find out all the direct paths from the SIP request sender to the user that the request is destined to. Afterwards, all the found paths will be rated and the RNM will take the average of the different ratings. For simplicity seek, the user who sent the SIP request will be denoted by destination and the user that the request is destined to, will be denoted by source. If the source and the destination users are directly linked, so no need to compute a reputation value as the destination identifier is already in the contact list of the source. If the source and the destination are not directly linked, the reputation rating will be the average of the reputation values for the destination returned by each of the neighbors of the source.

Defining the metrics to be used in the rating is always a hard task when building reputation systems. In general, reputation is computed locally, as in our case, and the obtained values need to be

combined in order to provide a global one. Combining these reputation values can be achieved in different ways that each of them reflects some specific purposes. In our context, we have chosen a simple metric that takes into account how trustful a recommendation "provider" is when his recommendation is being used. For instance, if we assume that p paths including nodes N_1,…, N_{n_p} exist between the source and the destination, the metric for computing the reputation of the destination node is described by the following formula,

$$r(source, destination) = \frac{\sum_{p}\left[r(source, N_1) * \prod_{i=1}^{n_p - 1} r(N_i, N_{i+1}) * r(N_{n_p}, destination) \right]}{p} \quad (1)$$

where r(A, B) is the reputation value expressing the amount of trust that A has in B. One can note that in the above formula, recommendations (i.e reputation values provided by the next nodes in the paths) are used according to the amount of trust that the user computing the reputation has in the recommenders. Based on the mentioned formula, the algorithm used for rating the incoming SIP request is depicted below,

```
Rep_eval(source, destination)
Start from the source
Mark the source as visited
If the source has no rating for the destination
α = 1, β = 0
for each u in the neighborhood of the source
    if u is not visited yet
        β ++
        α =  α * Rep_eval (source, u) * Rep_eval (u, destination)
        Mark u as visited
    Rep_eval (source, destination) = α/ β
Return Rep_eval(source, destination)
```

3.3.3.2 Spam threshold reputation

As mentioned earlier, this reputation value is needed to classify whether the request is spam or not. If the request reputation value is less than the threshold, the request is classified as spam, otherwise, it is further processed. The reputation threshold is tightly bound to the metric used for the rating, so if the metric changes, the threshold value might need to be updated in order to reflect more the context.

4 Further Work

In the future, we will consider more metrics and evaluate them in order to determine their accuracy. We will also study the impact of varying the threshold value on the different used metrics. The algorithm used for rating is particularly based on a "graph searching" procedure. As the SIP social network can be huge, searching for a path between the source and the destination may generate a

considerable delay. For this reason, we also have the intention to refine the mentioned algorithm and optimize it in order to cope more with large scale networks.

5 Conclusion

In this paper, we have presented an approach combining social networks and reputation rating to deal with SIP spam. After having described the problem space, we have defined a metric to be used in the rating operation as well as an algorithm designed for searching for eventual paths between the source and the destination and provide the corresponding reputation value. Our mechanism can work in concert with some other filtering techniques to prevent spam problems within SIP communities.

REFERENCES

[1] J. Rosenberg et al, "SIP: Session Initiation Protocol", RFC 3261, June 2002

[2] J. Rosenberg et al, "Session Initiation Protocol (SIP): Locating SIP Servers", RFC 3263, June 2002

[3] Earthlink, http://www.earthlink.net

[4] Iptel.org, http://www.iptel.org

[5] J. Rosenberg et al, "The Session Initiation protocol (SIP) and Spam", draft-rosenberg-sipping-spam-00, July 11, 2004

[6] J. Golbeck et al, "Trust Networks on the Semantic Web", http://www2003.org/cdrom/papers/poster/p314/p314_golbeck.html

[7] J. Golbeck et al, "Reputation Network Analysis for Email Filtering", http://www.ceas.cc/papers-2004/177.pdf

[8] P. Oscar et al, "Personal Email Networks: An Effective Anti-Spam Tool", http://arxiv.org/abs/cond-mat/0402143

[9] M. E. J. Newman et al, "Email Networks and the Spread of Computer Viruses", Physical Review E66, 0351001R (2002)

[10] M. E. J. Newman, "Properties of Highly Clustered Networks", Physical Review E68, 026121 (2003)

[11] M. E. J. Newman et al, "Finding and Evaluating Community Structure in Networks", Physical Review E69, 026113 (2003)

[12] J. M. Pujol et al, "Reputation Measures based on Social Networks metrics for Multi Agent Systems",

[13] D. Senn, "Reputation and Trust Management in Ad Hoc Networks with Misbehaving Nodes", Diploma Thesis DA-2003.27, July 2004

[14] A. Alexa, "Reputation Management in P2P Networks: The EigenTrust Algorithm", http://www.mpi-sb.mpg.de/units/ag5/teaching/ws03_04/p2p-data/01-20-paper2.pdf

[15] L. Xiong, L. Liu, "A Reputation-Based Trust Model for Peer-to-Peer eCommerce Communities", Proc of the IEEE International Conference on E-commerce (CEC'03), 2003

[16] B. Yu et al, "A Social Mechanism of Reputation Management in Electronic Communities", Proceedings of Fourth International Workshop on Cooperative Information Agents", pp 154-165, 2000

[17] G. Zacharias et al, "Trust Management Through Reputation Mechanisms", Applied Artificial Intelligence 14, pp 881-907, 2000