Analysis of Techniques for Protection against Spam over Internet Telephony

Vincent M. Quinten, Remco van de Meent, Aiko Pras

University of Twente, The Netherlands
v.m.quinten@student.utwente.nl,
{r.vandemeent, a.pras}@utwente.nl

Abstract. Spam in Internet telephony (SPIT) networks is likely to become a large problem in the future, as more and more people and companies switch from traditional telephone networks to Voice over IP (VoIP) networks, and as it is easy to spam VoIP users. The goal of this survey paper is to identify techniques to prevent and reduce SPIT. To compare the various SPIT protection techniques, criteria will be presented that must be met by these techniques. We also identify several combinations of techniques that complement each other, to increase the protection effectiveness.

1 Introduction

Spam over Internet telephony (SPIT) is defined as unsolicited bulk calls that result in media sessions, of which the content delivered to phone or voice terminals may include voice, images and / or video [1]. There are several kinds of SPIT, i.e. advertisement, telephone poll and telemarketing. Voice over Internet Protocol (VoIP) usage is growing fast; it is estimated that in the year 2010 25% of all households in Western Europe have abandoned traditional Public Switched Telephone Network (PSTN) services in favour of VoIP [2]. With the growth of VoIP communication, the 'abuse' of VoIP will grow as well. Advertisers who send numerous voicemail messages to VoIP users will cause a reduction in bandwidth and possible failures of the service, and are annoying for VoIP users. Compared to e-mail spam, the load on network resources by SPIT may be ten times as much [3]. SPIT is also more obtrusive [4], because the phone will ring with every spam message, even in the middle of the night, disturbing the users current activity. The use of VoIP instead of traditional PSTN networks will make it easier for spammers to make automated tools to deliver their spam to the user [1] and VoIP communication is also much cheaper, often flat-rate. The authors of [5] claim that the costs per call for VoIP are roughly three orders of magnitude lower than traditional PSTN calls, making it a lot cheaper for a spammer to get his message out into the world.

The two main protocols for VoIP are the H.323 protocol and the Session Initiation Protocol (SIP). Because of implementation errors and protocol features that may be exploited, both protocols will be equally vulnerable for SPIT [6]. Because of this, no distinction will be made between both protocols in the remainder of this paper.

Goal. This survey paper presents the state of the art of techniques to prevent or mitigate spam in VoIP networks. The paper discusses the various techniques and identifies

which combinations of techniques may be most promising for the future. This survey can be used as a reference to other researchers who want to develop new SPIT prevention techniques or improve existing techniques. Such paper was not available at the time of writing. However, two papers exist which discuss some related work. In [5] some techniques are discussed together with their advantages and disadvantages; that paper, however, only contains a small selection of spam protection techniques and aims at SIP techniques. A selection of the techniques described in the present paper is also described in [7], but [7] does not provide criteria upon which an analysis could be based.

Approach and organization. To provide this state of the art, a study of existing literature on the topic of SPIT has been conducted. Section 2 describes the criteria that SPIT prevention techniques have to meet; these criteria have been extracted from the literature. Section 3 lists the techniques, together with their main advantages and disadvantages. Section 4 provides an analysis of the effectiveness of these techniques, using the criteria from section 2. Finally section 5 contains some conclusions and remarks.

2 Criteria

For SPIT mitigation techniques to be effective and user friendly, they need to meet a number of criteria. This section discusses the most important criteria that have been distilled from literature on the topic of SPIT.

An important criterion is that protection techniques have to identify spam before the user's phone rings [1]. Because every time the phone rings it disturbs the user's current activity, spam is extremely annoying, particularly if the user wakes up in the middle of the night. A second criterion relates to maintenance of the protection technique. The less maintenance is needed by the user, the better [4]. So ideally the SPIT protection technique should be transparent to the user, which means that preferably it must be located at the service provider. The cost involved with spam protection is also an important aspect. Preferably the costs for the user are as low as possible, but the costs for the spammer should be as high as possible to assure that spamming becomes less profitable, which may eventually reduce the amount of spam [8]. The delay of the call caused by the protection technique is also an important factor to consider [4,8]. The less delay caused by the technique, the better; preferably the technique should not cause any delay at all, because long delays harm the direct nature of a telephone conversation. Another important issue is the impossibility to bypass spam blocking by spammers. Each spam protection technique will totally fail if spammers become able to circumvent the blocking [8]. A good spam protection technique should therefore be both effective as well as difficult to circumvent. Finally, the number of false positives and false negatives should be as small as possible, preferably even zero. This is important, because for many businesses and home users, telephone calls are very important. For businesses it's even essential that potential customers can reach the company [1].

3 Techniques

Based on the criteria defined in the previous section, this section discusses a number of techniques, including their advantages and disadvantages:

Signaling Protocol Analysis. VoIP calls consist of two parts: signaling and media data. Before a VoIP call starts, signaling data for setting up the call is exchanged between both users. Spammers are interested in the correct delivery of their calls, therefore the call routing information provided in the call setup request is valid and can therefore be used for further analysis. A characteristic of spam calls is that they are unidirectional: the spammer initiates the calls to the targeted network, but in general nobody calls the spammer. Another characteristic is the termination behavior; this is statistically consistent, so calls are generally terminated by the same party. A final distinction is that spammers in general do not call the same recipient for some period of time. Based on these characteristics, the authors of [1] defined a number of scenarios for termination behavior. Based on a statistical analysis of this behavior, the authors claim that it is possible to detect spammers with an accuracy of about 99.9%.

This technique has the benefits that it decides if a call is a spam call before the phone at the receiving side rings and the technique is located at the service provider, so the user isn't bothered with spam calls and maintenance. However this technique can only decide if a call is a spam call after at least ten calls from one caller. This indicates a heavy reliance on the fact that a spammer will not change his number for quite some time. But in reality it is quite easy to change your number in VoIP systems, this will make this technique relatively easy to circumvent. *Signaling protocol analysis* will also block some legitimate services, for example an automated system of the bookstore to inform you, your book has arrived. However this technique is fairly new, there is only one article [1] published about *signaling protocol analysis* to prevent spam in VoIP networks.

Do Not Call Registers are agreements between telemarketers that have agreed on that they will not call users that registered their phone number. Such a register can be controlled nationwide, so the user only has to register his phone number on one place and the control organization will handle the rest of the administration, so user maintenance is minimal [9]. This control organization also gives penalties to telemarketers who don't obey their agreement, this enforces the effectiveness of the register [10]. But a big disadvantage of *do not call registers* is, that it's an agreement of telemarketers, when a telemarketer has no agreement with the do not call register he can still spam registered users, this also holds for outsourced spam sources off-shore [1]. Off-shore spam sources are not very unlikely, because the costs for calling to another country are much less compared with the costs for normal PSTN calls. In contrast the costs involved with maintaining the do not call register and the investigation of complains can be high. These costs are eventually paid by the TAX payer instead of the spammer [10].

Circles of trust, as described in [5], is another technique that is very similar to a do not call register. Companies agree to exchange VoIP calls amongst each other and also agree to introduce a fine should one of them being caught spamming. Each company enacts measures to terminate employees who spam from their account. Circles of trust work well on small domains, but it is unknown how they would scale in large domains.

Whitelisting is a technique primarily used in instant messaging networks. In case of VoIP a whitelist contains the telephone numbers of the people that are allowed to call you, all other people are blocked. *Whitelisting* blocks all spam calls in theory, assuming nobody on your whitelist is a spam source or will become one [7]. But this is also a big disadvantage, because if a unknown person wants to call you, his call will be blocked.

Some home users don't think this is a disadvantage, but for business users it is vital that potential customers are able to contact them [7].

Whitelists are difficult to circumvent [5], because a change of identity will have no use. They also give the user complete control over who can and who can't call them, but this comes at a price. When users receive often calls from new callers, the amount of maintenance to the whitelist can be considerable.

Blacklisting is the complete opposite of *whitelisting*, instead of maintaining a list of numbers of the people that are allowed to call you, you maintain a list of numbers that aren't. For this system to be effective it needs to be implemented on a global level, when separate users maintain their one blacklists it will have very limited effects, because spammers will simply call someone else. Only on a global scale the costs for spammers can raise that much that spamming becomes unprofitable [7]. However everybody should be able to add a number to the blacklist, so a non-profit organization is needed to maintain the blacklist [7]. But even on a global scale it will not prevent users from receiving spam, because before a number is added to the blacklist a certain number of users have to answer the call and file a complaint and as mentioned before it's very easy to change your number in VoIP systems [5,7]. The use of proxy's in combination with blacklists also cause some unwanted side effects, because every user behind a proxy will be blocked when it is added to the blacklist, not just the spammer [5,7].

Greylisting applies a simple rule to all incoming calls: each call will be blocked, unless the same number has tried to establish a call within the last N hours/minutes. When a call is blocked, the sender will receive a message like "the user is currently busy". When the sender calls again within the N hour timeframe, his number is automatically added to a whitelist and all future calls will be connected immediately, requiring very little work from users [8]. *Greylisting* will not suffer from false positives when the used VoIP protocols are implemented correctly [7,8] and it will increase the costs for the spammer because he needs to call every user at least twice within the timeframe to make a successful call [7], which are great benefits.

According to [8], trying to circumvent *greylisting* has no effect or even an opposed effect, because *greylisting* will make other techniques even more effective, but we believe that spammers will adopt their systems to call twice within the timeframe. The delay caused by the rejection of the first call attempt makes *greylisting* unsuitable for most business users and emergency or other urgent calls [7].

A techniques similar to *greylisting* is described in [11]. Consent based communication, as described in [5], is also similar, except for automatically adding numbers to the whitelist, this causes an extra way to circumvent the technique by flooding the users with consent requests instead of spam calls [5].

Rate limiting allows the user to make a certain amount of calls per day. When the user exceeds the limit, he is likely to be a spam source and will be blocked. This simple technique doesn't bother average users, but will limit spammers [7]. But it has to be supported by all service providers worldwide, because otherwise spammers will just switch to a more spam-friendly provider [7].

Reputation filtering is a system adopted from instant messaging, where users can give each other reputation scores, based on this score users can decide to allow or reject a

call. When a call is allowed the number is added to a whitelist, comparable to buddy lists in instant messaging networks. However spammers are able to cheat in large networks by helping each other to receive a good reputation, which makes the system useless [5]. And the delay caused by reputation search paths can become very long in large networks which is not preferable [12]. As pointed out earlier with *whitelisting* user maintenance can become intensive [12].

Handshake/Challenge/Turing Tests, further referred to as *Turing tests*, is a technique adopted from e-mail and depends on the fact that some things are easy for humans, but almost impossible for computers [5]. However there is evidence that there are systems that can circumvent this technique, like described in [13]. In VoIP systems a user could answer a spoken question, for example a little math question and when the user provides the correct answer he is instantly connected [5]. Since speech recognition is difficult for computers with today's technology this will hold of automated spam calls. Added benefit in contrast with *greylisting* and *memory bound functions* the call will not lose its instant character [7]. But the system will not be very well suited for internationalization, because of the spoken question and the technique relies heavily on user acceptance making it not very attractive for business users [7].

Payments-at-risk tries to make spam more expensive for the spammer, but minimize the costs for the user. To achieve this, the calling party has to make a small deposit to the called party before the calling party can make the call, which will be refunded when the called party doesn't mark the call as being spam [5]. The problem with this system is that the costs for micro payment transactions make normal users lose money on every call. An example in [5] calculates that this will cost a normal user about $1.95 a month, when the user receives about 10 calls a day from unknown senders, which is relatively inexpensive.

Content Filtering makes use of speech recognition technology to analyze if the content of a message is spam, however with today's technology it's impossible to analyze the content real-time [5]. The system kicks in when the user has already answered the phone and is already disturbed by the spam call making the system ineffective, but it could be used to analyze voicemail messages. However VoIP providers that use some kind of encryption for extra security [14] or spammers trying to circumvent the system by using bad grammar or an accent [5] will provide even more difficulties for *content filtering* and probably even cause the technique to fail.

Memory Bound Functions. The basic idea of *memory bound functions* is: "If I don't know you and you want to send me a message, then you must prove that you spent, say, ten seconds of CPU time, just for me and just for this message" [15]. This "proof of effort" is mainly cryptographic, it's hard to compute but very easy to check. This "proof of effort" will consume computing power at the senders' device for every call, which implies that a spammer needs much more hardware for the same amount of calls and also has to pay the extra calling costs [7]. Average users will not be bothered by the small delay introduced according to [7] and normal calls will not be blocked, so it is still possible to make legal advertisement calls [7]. However for urgent calls and emergency calls the system is completely unacceptable [7].

4 Effectiveness Analysis

This section combines the results from the previous two sections and contains an effectiveness analysis for the techniques described before. The criterion of false positives and false negatives could not be considered, because at the time of writing no figures on this where available.

Unsuitable techniques. We identified *content filtering*, *do not call registers*, *reputation filtering*, *rate limiting* and *blacklisting* as being unsuitable techniques to prevent VoIP spam, because they don't meet the criteria defined in section 2. *Content filtering*, *blacklisting* and *rate-limiting* don't work before the phone rings, so they still allow disturbance of the users current activity. *Blacklisting* and *reputation filtering* additionally need a lot of maintenance when calls are received from many different sources. For *blacklisting* to be effective, implementation on a global scale is necessary, but this involves high costs for maintenance and control. *Do not call registers* also suffer from high maintenance costs; these costs will eventually be paid by the receiver of spam and not by the spammer. Reputation filters cause long delays when used in large networks, degrading the instant character of telephone communications.

Rate-limiting, *content filtering* and *blacklisting* are not feasible, because they respectively only limit the spammer in their ability, need technology that's not available and cause unwanted side-effects. Furthermore, all these techniques suffer from the fact that they are easy to circumvent.

Techniques with potential. The techniques that, in our opinion, show potential to become a suitable technique against SPIT or that will be effective in combination with other techniques are *whitelisting*, *signaling protocol analysis*, *payments at risk*, *greylisting* and *Turing tests*. The authors of [1] claim that *signaling protocol analysis* should have an effectiveness of about 99.9%; if this claim holds in practice, this would be a very good technique, but practical tests need to be done. A technique that's also very effective, but requires a lot of maintenance is *whitelisting*. *Whitelisting* is a qualified as a potential technique, because it shows some nice properties when used in combination with other techniques. *Greylisting* is a technique that needs to be used in combination with other techniques by design and it aims at raising the costs for spammers, because the spammer needs to call everyone at least twice. A main disadvantage of *greylisting* is the delay it causes, so it's not suitable for urgent or emergency calls. *Payments at risk* also raise the costs at the place it really hurts, at the spammer, thus make spamming less profitable and therefore reduce the amount of spam in the future. *Turing tests*, on the other hand, try to ensure a spammer has a hard time meeting the challenge presented to him; when a spammer succeeds to circumvent the *Turing tests*, it's easy to make the question more difficult. However, *Turing tests* require a certain amount of knowledge, which can be a problem when, for example, a child tries to call her father at work and is presented with the question "What is the capital of Italy?". Also, when the question becomes more difficult, the checking of the answer may become more difficult. If the buttons of the telephone are no longer sufficient to provide the answer, speech recognition may seem to be a solution. But, as described before, speech recognition is still unfeasible with today's technology.

Suitable techniques. The only technique that fulfills most criteria and is therefore in our opinion suitable, is *memory bound functions*. *Memory bound functions* work before the phone rings, increase the costs for spammers and are located at the service provider, thus are completely transparent to the end-user. The only criterion that isn't met by *memory bound functions*, is that it causes a small delay in all calls, but this can be solved easily, as will be described in the next section.

Combination of techniques. Using combinations of techniques can cancel out most of the disadvantages of the stand-alone techniques. The combinations we identified are *Turing tests* together with *whitelisting*, *memory bound functions* with *whitelisting* or *signaling protocol analysis* with *whitelisting*. In combination with the first two techniques, *whitelisting* reduces the delay, because the challenge or "proof of effort" needs to be done only once; afterwards the number of the caller is added to the whitelist and further challenges are skipped. This also partially solves the knowledge problem with *Turing tests*. A possible disadvantage of this combination, however, can be that spammers will adapt to the system [7], but this could be resolved by adding an expiration time to the entries on the whitelist.

Whitelists in combination with *signaling protocol analysis* reduces the amount of false positives [1], because automated services like the one of the bookstore described earlier could be added to the whitelist.

5 Concluding remarks

In this paper we have identified a number of protection mechanisms to prevent SPIT. These techniques can be divided into 4 categories: unsuitable techniques, techniques with potential, suitable techniques and combinations of techniques. In this paper the following techniques have been categorized as unsuitable: *content filtering*, *do not call registers*, *reputation filtering*, *rate limiting* and *blacklisting*. The following techniques have been categorized as having potential: *whitelisting*, *signaling protocol analysis*, *payments-at-risk*, *greylisting* and *Turing tests*. One technique is categorized as being suitable: *memory bound functions*. This technique fulfills almost all criteria defined in section 2 and is therefore suitable for use. *Turing tests* in combination with *whitelisting*, *memory bound functions* in combination with *whitelisting*, as well as *signaling protocol analysis* in combination with *whitelisting* are combinations that cancel out almost all disadvantages of each other.

There is not a single technique or combination of techniques that is the most promising approach for the future, but there are several options. Practical tests with these techniques need to be conducted to show which technique(s) will be best for use.

SPIT protection will probably always stay an arms race between spammers that try to circumvent protection techniques and researchers that develop new techniques or improve existing ones. However, as shown in this paper, by combining multiple VoIP spam protection techniques, many of the disadvantages can be canceled out.

We conclude with some remarks on possible future research. Most of the described techniques are easy to circumvent, just because it's easy to change your identity/phone number in VoIP networks. To prevent this kind of circumvention, research could be done to improve the system that allows someone to request new numbers. The effectiveness of

almost all VoIP spam prevention techniques would benefit from this. Measuring effectiveness is also a subject that needs further research, as for most described techniques no real test data is available regarding their effectiveness. To verify the various techniques, practical tests need to be conducted to draw better conclusions on which techniques are effective in practice. The *Turing test* technique also needs further research, the determine how to overcome the language barrier when the system is used in an international context.

References

1. R. MacIntosh and D. Vinokurov. Detection and mitigation of spam in IP telephony networks using signalling protocol analysis. pp. 49-52, 2005.
2. ElectricNews.Net. Mobile and VoIP to inherit the earth. http://www.theregister.co.uk/2005/06/27/rising_mobile_voip_revenues, (15-09-2006), The Register, 2005.
3. Ronald P. Gagner. Voice over Internet protocol: Secure or not recommendations to the business and private sector. Bowie State University, Maryland, 2005.
4. J. Pessage and J. Seedorf, "Voice over IP: Unsafe at any Bandwidth?," in Eurescom Summit Heidelberg, 2005.
5. J. Rosenberg and C. Jennings, "The Session Initiation Protocol (SIP) and SPAM," in Internet Draft, 2004.
6. E. Edelson. Voice over IP: security pitfalls. Network Security, vol. 2005: pp. 4-7, 2005.
7. Till Andreas Radermacher. Spam Prevention in Voice over IP Networks. University of Salzburg, Salzburg, 2005.
8. Evan Harris. The Next Step in the Spam Control War: Greylisting. http://projects.puremagic.com/greylisting/whitepaper.html, (24-09-2006), Evan Harris, 2003.
9. Federal Trade Commission. National Do Not Call Registry. http://www.ftc.gov/donotcall/, (14-09-2006), 2005.
10. Todd Edwards, "A Review of Southern States' No-Call Registries," in Southern Legislative Conference Atlanta, 2002.
11. N.J. Croft and M.S. Olivier. A Model for Spam Prevention in IP Telephony Networks using Anonymous Verifying Authorities. In ISSA 2005 New Knowledge Today Conference, South Africa, 2005.
12. Y. Rebahi and D. Sisalem. SIP Service Providers and The Spam Problem. In Voice over IP Security Workshop, Washington, 2005.
13. Sam Hocevar. PWNtcha - captcha decoder http://sam.zoy.org/pwntcha/, (20-12-2006), 2005.
14. Skype. Skype Help. support.skype.com, (23-11-2006), 2006.
15. Dwork Cynthia, Goldberg Andrew, and Naor Moni. On Memory-Bound Functions for Fighting Spam. 2003.