Using Cross-Media Relations to Identify Important Communication Requests: Testing the Concept and Implementation

Kumiko Ono
Columbia University
New York, USA
kumiko@cs.columbia.edu

Henning Schulzrinne
Columbia University
New York, USA
hgs@cs.columbia.edu

ABSTRACT

Important calls that originate from persons or organizations connected to the callee with weak social ties are often mistakenly labeled as unwanted ("spam", "SPIT") since their contact address is not found in the callee's address book. We have focused on the fact that the weak social ties are usually established through other communication means such as a web transaction for online purchase, travel reservation, or social media. We hypothesize that prior contact is a helpful distinguishing feature between important (or non-spam) calls and unwanted ones, and have proposed two mechanisms using cross-media relations such as *web-then-call and email-then-call*. One proposed mechanism is to help the callee collect addresses of potential callers in order to determine whether or not to answer a call using a caller ID. Another is to help the callee identify important calls using a proof of prior contact. We introduce an implementation of our proposed mechanisms called CURE (Controlling Unwanted REquests) system as proof of concept. We also conducted an initial survey of email messages as call detail records substitutes in order to help prove our hypothesis. This survey shows that 52 % of all incoming email messages are triggered by *web-then-email* relations, but carried unknown sender IDs. This result demonstrates that using cross-media relations can be potentially effective in identifying important calls and would be useful as an additional component of a spam filtering system.

Keywords

Unwanted calls; SPIT prevention; cross-media relations; email; WWW; web; VoIP; SIP

1. INTRODUCTION

Receiving unwanted bulk calls, e.g., telemarketing, charities, or polls, will likely be the same or worse problem than unwanted bulk email [1] due to the intrusive nature of calls, once large-scale open interconnected VoIP (Voice over IP) systems emerge. Caller-ID-based filtering systems [2, 3] have been widely used, but cannot always determine whether an incoming call is desirable to be answered. Figure 1 categorizes incoming calls into three groups according to their caller ID: non-spam calls determined by white-listing, unwanted calls determined by black-listing, and unlabeled calls, of which caller ID cannot determine whether or not to be spam, depicted on a gray background. Examples of these unlabeled calls include confirmation of appointments, travel reservations, or deliveries, and recorded notifications of school closing on a snowy day. These non-spam calls are often mistakenly labeled as unwanted since their caller IDs are not found in the callee's white list or address book. Our challenge, therefore, is to establish more sophisticated filtering mechanisms which can identify important calls that originate from persons or organizations whose contact addresses are not found on the callee's white list.

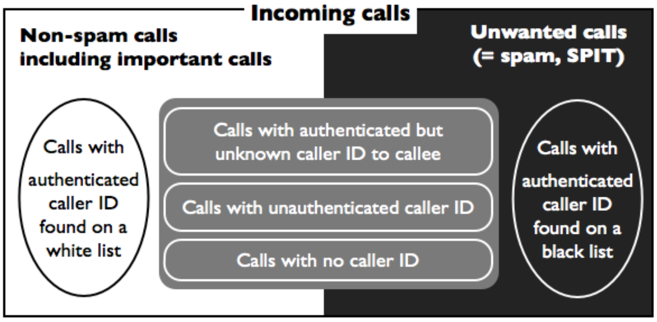
Incoming calls
Non-spam calls including important calls
Unwanted calls (= spam, SPIT)
Calls with authenticated caller ID found on a white list
Calls with authenticated but unknown caller ID to callee
Calls with unauthenticated caller ID
Calls with no caller ID
Calls with authenticated caller ID found on a black list

Figure 1: An overview of classification of incoming calls

We have observed that non-spam calls carrying unknown caller ID usually originate from persons or organizations who have had *prior contact* through a web transaction or email exchanges and have established weak social ties [4] to the callee. Another observation [5] shows that many callees now prefer receiving a call preceded by text message exchanges in order to avoid being disturbed even by a call carrying familiar caller ID. We, thus, have focused on this prior contact as an additional base of white-listing, rather than enhancing black-listing. In order to evade being blocked, VoIP callers can effortlessly pick a new caller ID, similar to an email address. This makes black-listing less effective for VoIP calls.

Our proposed mechanisms [6] use *cross-media relations* as proof of prior contact based on our hypothesis that the prior

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IPTComm 2011 Chicago USA
Copyright 2011 ACM X-XXXXX-XX-X/XX/XX ...$10.00.

Table 1: Cross-media relations – Types and information exchanged in prior contact

Types	Mechanism 1: Contact addresses of potential caller	Mechanism 2: Weak secret provided by user of CURE system (callee)
Web-then-call	Contact addresses either in plain text or hash format	Customized contact address of the callee
Email-then-call	Contact addresses	Message ID of an outgoing email message

contact is a helpful distinguishing feature between non-spam and spam calls. We describe cross-media relations further in Section 2. We then introduce an implementation of our proposed mechanisms using cross-media relations in Section 3, discussing the difficulties we encountered in making the specifications more detail. To test the concept of cross-media relations, we discuss our initial survey of email messages as call detail records (CDRs) substitutes in Section 4. Related work including the difference from our earlier effort [6] is described in Section 5. Finally, Section 6 concludes this paper.

2. OUR APPROACH USING CROSS-MEDIA RELATIONS

We hypothesize that a significant fraction of incoming calls are non-spam calls and originate from persons or organizations connected with weak social ties. Similar to spammers, their contact addresses are rarely found in the callee's address book. Those connected with weak social ties, however, differ from spammers in that they have had prior contact, directly or indirectly, with the callee through different communication means. Spammers usually make calls without any prior contact in order to make bulk calls efficiently. Thus, we hypothesize that prior contact is a helpful distinguishing feature between non-spam calls and unwanted ones.

Based on this hypothesis, we have proposed that both parties exchange an additional piece of information which can be used in future calls as an indication of prior contact [6]. Our proposal consists of two mechanisms depending on the type of information exchanges in prior contact. One is to collect contact addresses of potential callers. Another is to use a weak secret provided by the callee. Table 1 summarizes these two pieces of information varying according to the type of communication means in prior contact, namely the type of cross-media relations, *web-then-call* or *email-then-call*. These cross-media relations data will be described further in Section 3. Our proposal, thus, adds new filtering conditions using these cross-media relations into an existing call filtering system.

3. IMPLEMENTATION OF CURE SYSTEM

Figure 2 illustrates an overview of CURE (Controlling Unwanted REquests) system which we have implemented as proof of concept. Bob, a user of CURE system, is enabled to correlate an incoming call and prior contact with Alice, a potential caller, using cross-media relations data. The data are stored in a database, which allows multiple applications to update and query on behalf of the user, Bob. When updating the database, these applications add corresponding information such as the sender ID and destination addresses of an email message and the URL of a web site.

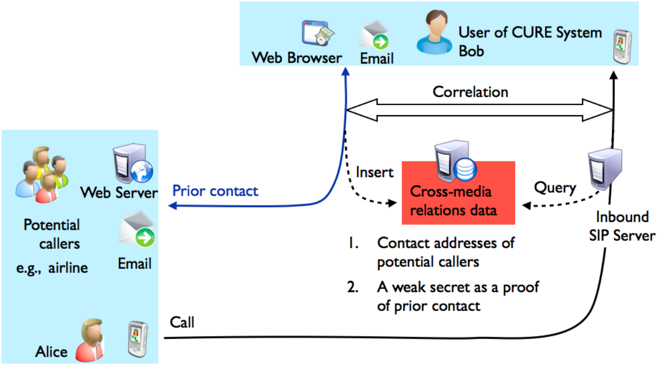

Figure 2: CURE system using cross-media relations

We use MySQL [7] database server running under Linux. We added RESTful JSON web interfaces [8] as database interfaces. To implement each of our proposed mechanisms in user applications, we have developed a Firefox Add-on [9] and a dedicated IMAP [10] mail client, and modified SIP communicator [11] as a SIP UA (User Agent). We use OpenSER [12] as an inbound SIP proxy server for the users of CURE system. We also use Apache [13] as a web server on the potential caller side.

The following subsections outline two mechanisms, collecting contact addresses of potential callers and using a weak secret, respectively in two scenarios: web-then-call and email-then-call. We finally describe the implementation of our proposed filtering system.

3.1 Collecting Contact Addresses of Potential Callers

Our first mechanism is to collect as many contact addresses of potential callers as possible. The more contact addresses of potential callers a callee can obtain beforehand, the more effectively existing caller-ID-based filtering systems can be used.

In addition to collect contact addresses in plain text, we have proposed a mechanism for collecting hashed contact addresses for web-then-call relations by considering privacy concern over user profiles at social networking web sites. Users in social network services prefer concealing their routable contact addresses on their published profile, but being identifiable at the callee. To fulfill their requirement, we have proposed that potential callers announce their contact addresses in hash format with privacy protection, namely contact addresses only for the purposes of identification. This will encourage users to publish contact addresses on their profile page.

3.1.1 Scenario 1: Web-then-Call Relations

We describe how to collect contact addresses of potential callers in two web-then-call scenarios: one collecting contact addresses in plain text from a web site, and another collecting hashed contact addresses of friends from a social networking web site.

Figure 3 illustrates an example of message exchanges in the first web-then-call scenario between Alice, a person from an airline service, and Bob, a user of the airline's web server. When Bob successfully signs up for a service, the airline's

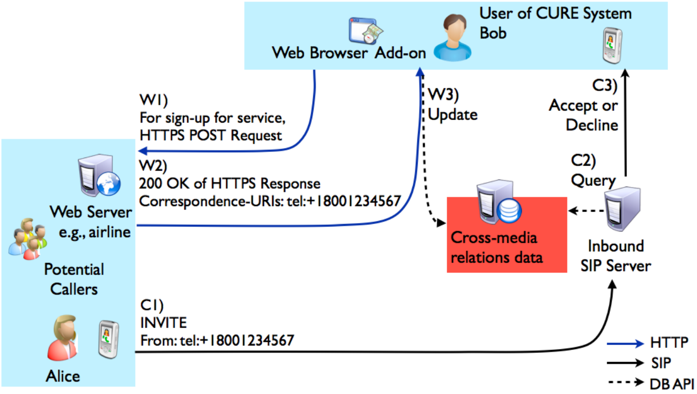

Figure 3: Collecting contact addresses of potential callers: Web-then-call

```
<html>
 <head>
   <META HTTP-EQUIV="Correspondence-URIs"
    CONTENT="email:customercare@example.com,
    sips:customercare@example.com,tel:+18001234567">
 </head>
</html>
```

Figure 4: An example of Correspondence-URIs in HTML

web server responds with an HTTP 200 OK response carrying their contact addresses in plain text for future communication. These contact addresses are conveyed in a new HTTP header, Correspondence-URIs [14] . This transaction should be protected by using secure HTTP (HTTPS) [15]. A web server application needs to generate an HTML data including contact addresses in Correspondence-URIs in an HTML META tag, HTTP-EQUIV [16], as depicted in Figure 4. The web server then converts the HTTP-EQUIV tag into the corresponding HTTP header.

When Bob receives the HTTP 200 OK response, the Firefox Add-on extracts the contact addresses in the Correspondence-URIs header field, and updates the database of cross-media relations with the URL of the web site for reference. To prevent misuse, the Add-on prompts Bob for confirmation before the update. Additionally, the Add-on asks Bob if the expiry date should be set for the temporal use of the service and when it will expire.

If Alice needs to contact Bob regarding his signed-up service afterwards, she just needs to make a call to him from one of the contact addresses which have been delivered in the previous web transaction. An inbound SIP server for Bob queries the database for the caller ID in order to determine whether to accept the call.

Next, we describe the second web-then-call scenario, where contact addresses are in hash format, using social network services. We assume that Bob, a user of social network services, can retrieve hashed contact addresses of his friends including Alice through a JSON object in a HTTP 200 OK response to an HTTP GET request. The hash string is generated by an application at a social network web site using SHA-1 [17] from a contact address in plain text concatenated with the URL where the address is published. This concatenation is to prevent hashed contact addresses from being correlated across different web sites.

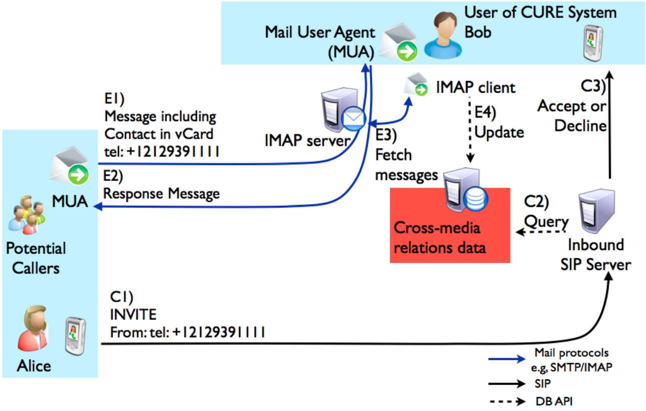

Figure 5: Collecting addresses of potential callers: Email-then-call

The Firefox Add-on then extracts these hashed contact addresses in the JSON object and updates the database with the site's URL. This URL is used for verifying hashed contact addresses in an incoming call at a call filtering system. In the same way as the first web-then-call scenario, the Add-on prompts Bob for confirmation and preferred expiry date before updating.

When Alice makes a call to Bob, she needs to specify her hashed contact address published on her profile page in order to be identified as a person in Bob's social graphs. The modified SIP communicator set a new SIP header, Sender-Ref [18], to the hashed contact addresses with the h-contact type parameter in addition to the originator's contact address in From header. An inbound SIP server for Bob queries the database for the Sender-Ref header field in addition to the caller ID, in order to verify the hashed contact address and determine whether to accept or decline the call.

3.1.2 *Scenario 2: Email-then-Call Relations*

Figure 5 illustrates that Bob collects Alice's contact addresses in a vCard [19] attachment to her email message only when he replied. Checking his reply is important to limit contact addresses to those from non-spam messages since the reply, except from a compromised machine, is evidence that the original message was a non-spam.

We have implemented an IMAP client dedicated to our proposed mechanisms, instead of adding required functions into an existing mail client. This is because required functions need to run periodically without any user interaction. This IMAP client first fetches non-spam incoming messages by determining whether or not to be replied and extracts contact addresses in a vCard. The IMAP client then inserts the contact addresses into the database.

How to determine whether a message was replied was not so easy and quick as we had expected. We had expected that we could identify replied messages by checking the \answered flag. However, sending or storing the flag is an optional operation according to the IMAP specification. In fact, Thunderbird mailer 3.1 [20] does not send the flag nor Gmail [21] IMAP server and its web mail client do not store the flag. In these cases, instead of checking the flag, the IMAP client needs to look up for saved replies. This operation requires higher cost in processing than the flag-based operation, but the computational cost of accumulating the recipients' mail addresses of outgoing messages are not expensive and widely used for automatically generating an address book or white-list. We, thus, implemented both two ways to identify replied messages on our IMAP client.

3.2 Using a Weak Secret

Another type of cross-media relations data is a weak secret provided by the callee. This weak secret is confidential information which is strong enough for determining whether to answer a call, which is a relatively low-risk interaction.

3.2.1 Scenario 1: Web-then-Call Relations

A web transaction, namely an HTTP transaction, does not have any transaction ID according to the HTTP specification [22]. We, thus, have proposed using a customized contact address containing a random component as a proof of a web transaction.

When Bob fills out contact information in a sign-up form, the Firefox Add-on we implemented helps him generate a random component by clicking a button and set between the user name and the domain name preceded with +, in the same way as the email addressing practice called "subaddressing" [23]. For tel-URI [24], random digits follow the E.164 number like an extension.

When Alice makes a call to Bob, she can be identified by the destination address (e.g., `sip:bob+anwxv3fa@columbia.edu`), not by her caller ID. Although no extension is needed in a SIP client, the SIP server, OpenSER, had to support subaddressing of the destination address in its routing process. A SIP server then queries the database for the destination address to determine whether to answer or decline the call.

3.2.2 Scenario 2: Email-then-Call Relations

Unlike an HTTP transaction, an email message has a global unique ID in the Message-ID header field [25], which is generated by a mail server or client.

Bob first sends an email message to a potential caller. The message ID of his outgoing message can be used as a weak secret to prove his acceptance for future communication. When Alice makes a call to Bob, she should specify the message ID of his message using the modified SIP communicator so that a SIP INVITE request conveys the message ID in the Sender-Ref header field. The SIP server, therefore, can determine whether to accept or decline the call using the message ID, regardless of the caller ID.

In addition to modifying the SIP communicator, we have implemented another dedicated IMAP client to periodically collect the message IDs of outgoing messages. To convey the message IDs securely, users should use TLS (Transport Layer Security) [26] for all the hops from a client to the other. However, using TLS alone is not secure enough, precisely not privacy-aware. We have a privacy concern over messages posted to a mailing list. This privacy problem is caused by the fact that many mailing lists publish their archives including their message IDs on the Internet. Thus, these message IDs of mailing lists in public should not be used as a weak secret.

How can we determine whether the destination of an outgoing message is a mailing list? If senders have explicitly subscribed to a mailing list, they can easily determine that. Some mailing lists, however, register subscribers and make their messages shared or public without their consent. Thus, senders cannot always know the destination is a mailing list. Even though senders cannot distinguish a mailing list address when sending a message, they can find that by checking List-* or Precedence headers when receiving a message from a mailing list. The IMAP client, therefore, needs to determines the destination address of a mailing list by received messages from the mailing list besides user's configuration. The IMAP client then excludes an outgoing email message destined to the mailing list when extracting the Message-ID header field, or deletes stored the message ID of the message destined to a mailing list if any. Since the sender subscribes to the mailing list, the sender can receive the message immediately after sending. Thus, this delay does not cause any problem. In this way, the IMAP client avoids using a published message ID as a weak secret.

3.3 Call Filtering

A call filtering system using cross-media relations can be located in an inbound SIP proxy server or a SIP UAS (User Agent Server). When receiving an incoming call, the filtering system reads the originator address in the From header authenticated using the Identity header [27], the destination address in the To headers, and the referred message ID in the email type parameter and hashed contact address in the h-contact type parameter of the Sender-Ref header. It then looks up these header fields on the access control list of the callee, the related URL, and the expiry date in order to determine whether to accept or decline the call.

We have implemented this call filtering system on OpenSER SIP server which queries the database for an access control list including cross-media relations, instead of using CPL (Call Processing Language) [28] script. Although the CPL script, which is written in XML [29], is a more general mechanism and enables users to control their services by themselves, it does not benefit the users of the CURE system. This is because the applications such as the Firefox Add-on, instead of users' direct control, update their CPL script on behalf of users. On the contrary, the CPL incurs the significant processing cost and storage cost of large-sized XML data. The SIP server, therefore, simply uses an access control list in MySQL database, in order to avoid a longer setup delay caused by the XML processing for the CPL.

4. TESTING THE CONCEPT BY OBSERVING EMAIL

Ideally, we would like to evaluate our implementation by putting it to trial, but three practical problems have hindered the trial. First, we currently receive very low volume of unwanted calls compared to unwanted emails. Second, the trial requires cooperation with web sites which operate call center services or social network services. Third, it requires end-to-end SIP connections to transfer a weak secret in a SIP message between a caller and the callee since some PSTNs (Public Switched Telephone Networks), for example in the US, does not allow to transfer the extension in the destination phone number at once. Although SIP has been used many real services, the majority of VoIP calls occur in interconnecting with PSTNs so far. Thus, instead of evaluating our implementation, we test the concept of cross-media relations by conducting a survey of incoming email messages as a substitute for incoming calls.

This survey measured the fraction of email messages triggered by cross-media relations, *web-then-email* and *email-then-email*, respectively, as an indicator of the potential effectiveness using web-then-call and email-then-call. The survey also measured the fraction of email messages triggered by *call-to-email* for reference.

We categorized incoming messages for 12 email accounts of our colleagues: five of university email account and seven of free email accounts like Gmail. We collected 7575 messages received for four weeks and saved in the participants mail boxes on IMAP servers; 3618 messages for their university email accounts and 3957 messages for their free email accounts. Note that these messages included only part of unwanted messages which the participants received since many unwanted messages had been removed by the spam filters on these IMAP servers, by the participant's mail client, and by the participants manually.

4.1 Classification of Incoming Email Messages

Figure 6 illustrates how each participant categorized incoming messages into 12 groups: 11 groups of non-spam messages and a group of unwanted messages. The first six groups are automatically determined as non-spam by an analyzer we developed. The "sent before" and "replied within two weeks" groups are determined with outgoing email messages. The "trusted users" and "trusted domain" groups are determined with the user configurations of the analyzer. Although the analyzer does not use participant's address book, these four groups are equivalent to non-spam messages determined by their address book. The next two groups, "FoF in CC" and "FoF in ML", are categories for messages from friends of a friend (FoF) determined by CC addresses of messages sent from trusted users and messages from other subscribers of the mailing lists (MLs) that the participants join, respectively. Thus, these two groups are determined by extended social graphs automatically calculated by email communication history.

After the analyzer automatically determined non-spam messages, the participants manually labeled messages according to cross-media relations to prior contact as "web", "email", and "calls". These three groups also contain non-spam messages, which are prone to be labeled as spam by conventional caller-ID-based filtering systems. These three groups, therefore, indicate potential effectiveness of using cross-media relations.

Yet, non-spam messages remain. The participants determined messages triggered by their publicity if they are well-known in a specific field. They then left remaining messages as "unlabeled" if the participants have no idea what triggered the messages. These two groups, therefore, indicate the limitation of our proposed filtering systems using cross-media relations.

4.2 Results: Fractions of Messages

Figure 7 illustrates the average percentages of incoming messages in these 12 groups. Each group has two bars: a dark (or blue) bar for university email accounts and a light (or green) bar for free email accounts.

The highlight of our results is that prior web transactions triggered 32.6 % of messages for university email accounts, 66.5 % for free email accounts, and on average 52 % for all email accounts. These results demonstrate that web-then-email relations will be very effective in determining non-spam messages while the other two types of cross-media re-

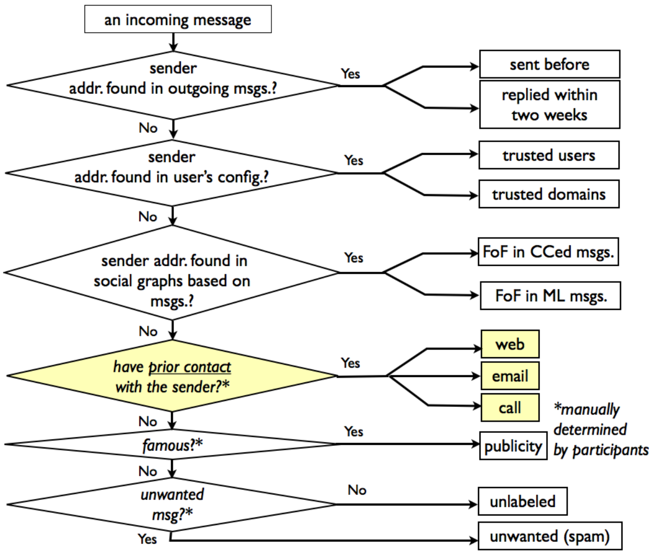

Figure 6: Incoming messages categorization flow chart

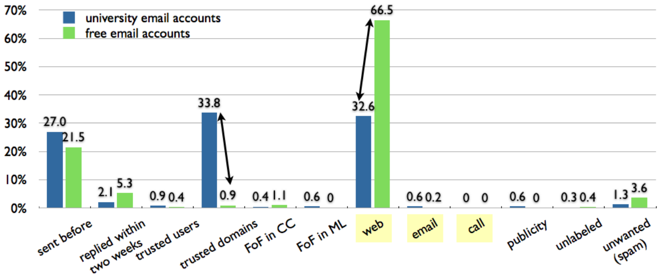

Figure 7: Percentages of incoming messages in 12 groups

lations, email-then-email and call-then-email, are scarcely used.

Although for most groups, fractions of messages are similar between these two types of email accounts, significant differences are found in two groups of messages: messages from trusted domains and messages triggered by web-then-email relations. Messages from trusted domains occupy 33.8 % of all incoming messages for university email accounts while they are only 0.9 % for free email accounts. This is because university email accounts, unlike free email accounts, are used for internal communication within university domains and also within professional communities. For the difference in messages related to web, we suspect that this is also caused by the difference in the usage of email accounts. Compared to university email accounts, free email accounts tend to be more easily given to the subscribers of on-line services and to receive many newsletter or purchase confirmations. These contrasts, therefore, indicate that the effectiveness of filtering systems using cross-media relations highly depends on the usage of user accounts.

Yet after using cross-media relations, a small fraction of non-spam messages remain. 0.6 % of messages for university email accounts are triggered by publicity and 0.3- 0.4 % of messages remain unlabeled. These indicate the limitation of our proposed filtering mechanisms. The fraction of messages in the publicity group would be larger if participants include professors. We have observed that these messages are often from other less-known members belonging to the same professional community. Thus, it would be helpful in identifying valid messages if we have a query mechanism about sender references, especially membership in a professional organization like ACM. This is left for our future work.

To find out a distinguishing feature between unwanted and unlabeled messages, we examined the "unlabeled" and "spam" messages in detail of the usage of subaddressing and hidden destination addresses by using BCC (Blind Carbon Copy) header, which is usually used for concealing other destination. However, we failed to find any significant characteristics in these usage. For both groups of unwanted and unlabeled, most messages were BCCed and subaddressing were not found.

In summary, although we still have a small fraction of unlabeled non-spam messages, we can conclude that using cross-media relations, especially to prior web transactions, would be effective in helping to label non-spam incoming email messages. As well as web-then-email, web-then-call relations would be effective. In addition to using cross-media relations, a query mechanism about sender references like membership could be another helpful component in identifying non-spam messages.

5. RELATED WORK

Similar to email spam prevention, there is no panacea for preventing unwanted calls; thus, a collection of solutions is needed. Conversely, unlike email spam prevention, we need to rely on the signaling message rather than content since content-based filtering cannot prevent call spam from disturbing callees since the signaling message has already interrupted them before content is sent. To prevent call spam, we have introduced a new approach by identifying desirable calls, namely a white-listing approach, instead of detecting undesirable. We, thus, discuss other white-listing approaches here.

As described in [30], the solution space can be divided into two categories: one places procedural, computational, financial, and/or legal burdens on the caller side, and another filters incoming calls on the callee side. Our proposed mechanisms using cross-media relations work in conjunction with adding procedural burden on the caller side and enhancing the filtering mechanism based on signaling messages on the callee side. Consent-based solutions in SIP [31] similarly approach to identifying non-spam message based on signaling messages. To grant a permission, whereas the consent-based solutions require additional SIP messages, our mechanisms use the existing messages without requiring additional round trip times. Instead, our mechanisms add a piece of information into the messages including communication means other than the SIP.

Most well-known solutions for labeling incoming calls are based on authenticated caller IDs. For VoIP calls using SIP, caller ID authentication requires the SIP Identity header, which is signed by domain similar to DKIM signatures [32] for email messages.

To maintain caller IDs in a white list updated, many approaches to augmenting contact addresses based on social graphs have been proposed. To expand white lists using social networks, Ceglowski and Schachter [33] introduced privacy-aware address book sharing by exchanging it as an email attachment while we [34] offered address book propagation within SIP messages. To update white lists based on communication history, Balasubramaniyan and his colleagues [35] introduced call credentials generated from the call history of a caller. Although their communication history is limited to calls, Shacham and Schulzrinne [14] addressed using alternative communication channels, namely using web transactions to collect contact addresses from potential callers. This is the base work for our first mechanism using cross-media relations, which we expand to use email exchanges.

To label incoming calls without caller IDs, one of our labeling mechanisms is based on the destination address with subaddressing [23], which has already been deployed for email messages. For calls, subaddressing of the SIP-URI is new, but the concept of extensions in the tel-URI is similar to call distribution at a PBX (Private Branch eXchange).

Our earlier effort [6] presented the concept of using cross-media relations motivated by the results of a preliminary survey of important calls carrying unknown caller ID. It also described an implementation to have for our proposed mechanisms. Based on this effort, we discuss the implementation which we have developed in details and show the results of email observation to test the concept in this paper.

6. CONCLUSION AND FUTURE WORK

Our implementation of CURE system demonstrated how we can use cross-media relations as a proof of prior contact in order to determine whether an incoming call is likely to be desirable to be answered. To test potential effectiveness of our proposed mechanisms, we observed by an initial survey that 52 % of messages can be identified as non-spam by using relations to prior web transactions. This result, thus, indicates that cross-media relations would be useful as an addition component of a call filtering system.

To provide mode evidence of the potential effectiveness, we plan to conduct the survey with more participants. To evaluate the practical effectiveness of using cross-media relations, we plan to test our concept in an integrated service with multiple communication means like Google voice [36]. We also plan to propose a query mechanism about sender or caller membership in order to enhance the capability of identifying important or non-spam communication requests.

7. ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Grant No. CNS-0905406. The authors thank our colleagues, especially Jan Janak for his help in modifying configurations of OpenSER. Yifan Zhang helped with the CURE system implementation. The authors also thank anonymous reviewers for their helpful comments.

8. REFERENCES

[1] L. Andersson, E. Davies, and L. Zhang. Report from the IAB workshop on Unwanted Traffic March 9-10, 2006. RFC 4948, IETF, August 2007.

[2] 3Com Corporation. 3Com V7000 IP Communications Platforms: Call restrictions.

[3] Skype Limited. Staying safe on Skype: Setting your Skype privacy levels. http://www.skype.com/intl/en-us/security/safety/staying-safe.

[4] M.S. Granovetter. The Strength of Weak Ties. *Amer. J. of Sociology*, 78:1360–80, May 1973.

[5] P. Paul. Cultural Studies: Don't Call Me, I Won't Call You. *The New York Times*, March 2011. http://www.nytimes.com/2011/03/20/fashion/20Cultural.html.

[6] K. Ono and H. Schulzrinne. Have I Met You Before? Using Cross-Media Relations to Reduce SPIT. In *ACM IPTComm*, July 2009.

[7] MySQL. http://www.mysql.com/.

[8] P. James. RESTful interface to MySQL using PHP. http://sourceforge.net/projects/phprestsql/.

[9] Add-ons for Firefox. http://addons.mozilla.org/.

[10] M. Crispin. INTERNET MESSAGE ACCESS PROTOCOL - VERSION 4rev1. RFC 3501, IETF, March 2003.

[11] SIP Communicator. http://sip-communicator.org/.

[12] Kamailio (OpenSER) SIP Server. http://www.kamailio.org/w/.

[13] Apache HTTP Server Version 2.0. http://httpd.apache.org/docs/2.0/.

[14] R. Shacham and H. Schulzrinne. HTTP Header for Future Correspondence Addresses. Internet-draft, IETF, May 2007. Work in Progress.

[15] E. Rescorla. HTTP Over TLS. RFC 2818, IETF, May 2000.

[16] D. Raggett, A.L. Hors, and I. Jacobs. HTML 4.01 Specification. December 1999.

[17] Secure Hash Standard. Federal Information Processing Standard (FIPS) 180-2, National Institute of Science and Technology, August 2002.

[18] K. Ono and H. Schulzrinne. Referencing Earlier Communications in SIP Requests. Internet-Draft, IETF, October 2009. Work in Progress.

[19] F. Dawson and T. Howes. vCard MIME Directory Profile. RFC 2426, IETF, September 1998.

[20] Mozilla messaging: Thunderbird. http://www.mozillamessaging.com/.

[21] Google. Gmail. http://mail.google.com/.

[22] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee. Hypertext Transfer Protocol – HTTP/1.1. RFC 2616, IETF, June 1999.

[23] K. Murchison. Sieve Email Filtering: Subaddress Extension. RFC 5233, IETF, January 2008.

[24] H. Schulzrinne. The tel URI for Telephone Numbers. RFC 3966, IETF, December 2004.

[25] P. Resnick. Internet Message Format. RFC 5322, IETF, October 2008.

[26] T. Dierks and E. Rescorla. The Transport Layer Security (TLS) Protocol Version 1.2. RFC 5246, IETF, August 2008.

[27] J. Peterson and C. Jennings. Enhancements for Authenticated Identity Management in the Session Initiation Protocol (SIP). RFC 4474, IETF, August 2006.

[28] J. Lennox, X. Wu, and H. Schulzrinne. Call Processing Language (CPL): A Language for User Control of Internet Telephony Services. RFC 3880, IETF, October 2004.

[29] T. Bray, J. Paoli, C. M. Sperberg-McQueen, E. Maler, and F. Yergeau. Extensible Markup Language (XML) 1.0 (Third Edition). W3C Recommendation REC-xml-20040204, World Wide Web Consortium (W3C), February 2004.

[30] J. Rosenberg and C. Jennings. The Session Initiation Protocol (SIP) and Spam. RFC 5039, IETF, January 2008.

[31] J. Rosenberg, G. Camarillo, and D. Willis. A Framework for Consent-Based Communications in the Session Initiation Protocol (SIP). RFC 5360, IETF, October 2008.

[32] E. Allman, J. Callas, M. Delany, M. Libbey, J. Fenton, and M. Thomas. DomainKeys Identified Mail (DKIM) Signatures. RFC 4871, IETF, May 2007.

[33] M. Ceglowski and J. Schachter. LOAF. http://loaf.cantbedone.org, 2004.

[34] K. Ono and H. Schulzrinne. Trust Path Discovery. Internet-draft, IETF, June 2006. Work in Progress.

[35] V. Balasubramaniyan, M. Ahamad, and H. Park. CallRank: Combating SPIT Using Call Duration, Social Networks and Global Reputation. In *CEAS 2007 - Fourth Conference on Email and Anti-Spam*, August 2007.

[36] Google. Google voice. http://www.google.com/voice.