

PHEmail: Designing a Privacy Honoring Email System

David H. Nguyen and Khai N. Truong
College of Computing & GVU Center
Georgia Institute of Technology
Atlanta, GA 30332-0280 USA
[dnguyen, khai]@cc.gatech.edu

ABSTRACT

Controlling one's personal and private information could help alleviate one of the greatest harms facing the Internet today – the loss of attention due to the over abundance of unsolicited email (spam). If one could control the dissemination and usage of one's email address, one could eliminate spam. We introduce a privacy honoring email system that leverages the user's social network to provide access control to the user's email.

Keywords

Privacy, email application, design, social network, spam.

INTRODUCTION

We take *privacy* to mean the right to control one's personal and private information. Controlling one's personal and private information could help alleviate one of the greatest harms facing the Internet today – the loss of attention due to the abundance of unsolicited email (spam) [2]. Because one does not have control of the distribution and dissemination of one's email address, that address could be used by anyone on the Internet. With the cost of sending email only negligible, spammers blanket the Internet with advertisements for ink cartridges, porn, cruises, weight loss programs, or whatever they might want to sell. As a consequence, people spend large amounts of their attention separating actual email and deleting spam.

Thus far, there are myriads of technical approaches to counteract spam. These schemes include black lists to restrict individuals to entire domains; white lists to allow selected individuals, one's address book is a common white list approach; Bayesian networks and other artificial intelligence techniques to categorize spam from valid emails; authentication of users and SMTP servers; the list of technical solutions continues. Economic solutions have recently been proposed. Kraut, et al. suggests adding a financial cost to sending email [2]. There are, of course, proposals for legal solutions. However, since this is a gray area (which includes arguments such as one person's spam is another's useful information and the

right to free speech), passing laws might prove to be very difficult. In addition, laws are bounded by borders, while the Internet is not necessarily so bounded.

Now, if we consider Lessig's model of regulation through architecture, social norms, market, and law, we can easily see that a missing approach is the social approach [3]. To that end, we propose Privacy Honoring Email (PHEmail), a system designed primarily on people's social norms and an understanding that the architecture of email and the Internet can be changed through code. We examine and enhance the social norms surrounding the use practices of email as a tool in human-to-human communication.

DESIGN

As people interact with each other, they leave traces of their social networks. These traces can be found in the people they work, play, and interact with, in the people they talk with, in the people they email [1,4], in the people they instant message (buddy list). And as more interactions are computer mediated, more social network traces can be collected and analyzed. These traces build little by little, but at the end, taken as a whole, the social network can be quite apparent.

We leverage this social network as a means for access control. Because of the type of private information an email address is, the problem need not be about who HAS the information (email addresses), but rather, who can USE that information. If filters worked well, people will not care who has their email addresses. Anyone can send them email, the filters will figure out who gets through. When we use a person's social network to further refine the filtering of email, messages from someone in that person's social network would be percolated up and be tagged as more important. "Regular" email (from someone outside the social network) would be effectively unimportant and be pushed down in the pile for later perusal (if at all).

As seen in research such as ContactMaps [4], a large part of one's social network can be gathered by going through one's email archive. A person to whom one sends and receives email regularly is highly likely a part of one's social network. The technique of counting one's email archive can identify a large part of one's social network. However, there are always exceptions and limitations. For example, a family member who does not email regularly is still a part of one's social network. Another

limitation of this quick technique is that there is no mechanism for social introductions. Therefore, the social network formed is only one layer deep. One can make assumptions about people's connections to each other through their associations of email addresses in the email headers [1]. However, the exact nature of the relationships still needs annotation.

We address these limitations by adding *tokens* to email messages.¹ While it is possible to completely reinvent email systems, we do not want to disrupt people's current practices. Therefore, the clients can essentially remain the same; it is the underlying servers (SMTP for sending and IMAP for receiving) that we are augmenting. We augment the servers by adding a proxy that handles the communication between the servers and the email client. The proxy automatically manages the transfer of PHEmail tokens for people using our system. For people not using our system, these tokens are meaningless data that their email clients will ignore. The tokens hold, among other things, information about the owner of the token, the holder of the token, the relationship between the holder and the owner, and a usage counter of how many times this token has been used. Of note, the token is encrypted and can only be decrypted, viewed, and changed by its owner.

How It Works

As we stated earlier, the tokens used in our system are encrypted. Moreover, the system is closed so only the intended recipient can view the tokens and tokens cannot be forged. To better help the reader to understand the system, we will describe a few scenarios of how the system would be used.

How dynamic privacy is honored

Privacy is very dynamic. Not everyone treats privacy the same way. Some are very open with their information; others are very close with the same information. Therefore, we wanted a design that allowed people to dynamically set their own privacy level. For email, that means allowing people to filter as much or as little email as they want. We want to give mechanisms to people, so they can put into effect their own privacy policies. We want to empower people, but also be mindful of defaults. So by default, we consciously decided:

1. To lightly filter, so all email will still get through, but with PHEmail messages flagged. The mechanism is there if users want to tighten the filtering.
2. To attach all outgoing emails with the owner's token. Because only the owners of the tokens can view and manage their own tokens, holding a token does not

necessarily mean having access to the owner. So to save face, tokens can be given out without sacrificing one's privacy.

Seeding the system

Bob Metcalfe is accredited for saying, "The value of a network increases exponentially with the number of nodes." In the same vein, the added value and effectiveness of the system will increase when more and more people use Privacy Honoring Email. However, as most of a person's social network can already be found in that person's email archive, leveraging the person's social network can start right away.

We begin seeding the system by counting a person's email archive (if the person keeps such an archive). Additionally, as users of the PHEmail system send each other email, they will automatically be exchanging tokens and annotating the social traces PHEmail uses.

Social introductions

When a person (A) wants to talk with someone (C) she does not know, often she seeks out a common friend (B) who would introduce her to this person (C). Using PHEmail, person A asks person B for person C's token. Then person A can email person C using the given token. When person C receives this email, the system will know that this email is from a person A, who is connected to person B, who is in person C's social network. Using the information stored in the token, person C can have filters such as: "give higher priority to messages from people less than three degrees of separation from me."

CONCLUSION

Privacy is a large problem. We have addressed the issue of spam as a privacy issue because of the loss of control of one's personal information (in this case, one's email address) leading to personal harm (loss of attention and time). In this paper, we have presented PHEmail to address this problem. We do not think PHEmail is the only needed solution in stopping spam. However, when used with existing solutions, it may provide a simple, but necessary component in ending spam. We will deploy the system to evaluate this hypothesis in a long-term study.

REFERENCES

1. danah boyd. *Faceted Id/entity: Managing Representation in a Digital World*. Cambridge, MA: MIT Master's Thesis. August 9, 2002.
2. Robert E. Kraut, James Morris, Rahul Telang, Darrin Filer, Matt Cronin, and Shyam Sunder. *Markets for attention: will postage for email help?* In Proceedings of CSCW 2002. New Orleans, LA.
3. Lawrence Lessig. *Code and Other Laws of Cyberspace*. Basic Books. 2000
4. Steve Whittaker, Quentin Jones, Loren Terveen. *Contact Management: Identifying Contacts to Support Long-term Communication*. In Proceedings of CSCW 2002. New Orleans, LA.

¹ While the initial design did use tokens, passed as MIME attachments, the actual current implementation of the system uses public/private keys to address security concerns. We retain the use of the term *tokens* for clarity in explaining the design of PHEmail.