A Comparison of PGP, IBE, and Password-based Secure Email

Scott Ruoti
Brigham Young University
Sandia National Laboratories
ruoti@isrl.byu.edu

Daniel Zappala
Brigham Young University
zappala@cs.byu.edu

Jeff Andersen
Brigham Young University
andersen@isrl.byu.edu

Kent Seamons
Brigham Young University
seamons@cs.byu.edu

Tyler Monson
Brigham Young University
monson@isrl.byu.edu

Abstract
In this paper, we describe a user study we are conducting that compares PGP, IBE, and password-based secure email. This will be the first study that compares wholly different key management schemes while holding the overall system and interface design constant, thereby reducing confounding factors. Our goal is to establish the inherent usability differences between PGP, IBE, and password-based encryption in secure email.

Author Keywords
Secure email, user study, key management

ACM Classification Keywords
H.1.2 [Information Systems]: User/Machine Systems—*Human Factors*; K.6.m [Computing Milieux]: Miscellaneous—*Security*

Introduction
Usable, secure email is still an open problem more than 15 years after it was first studied by Whitten and Tygar [15]. Even though recent research has made progress towards usable, secure email, it is still unclear how the usability of different key management schemes compares. While past studies have examined different key management schemes, the results of these studies cannot be directly compared as there are too many confounding factors; namely, the sys-

MessageGuard

MessageGuard is a research framework for rapidly building and comparing content-based encryption prototypes [9]. MessageGuard had several several benefits:

- Prototypes built using MessageGuard share the same basic system and interface design.
- MessageGuard's email interface has consistently been rated as highly usable [11, 7, 8].
- MessageGuard support pluggable key management, allowing us to easily switch out the key management scheme of each variant.

tems in which the key management schemes were implemented had entirely different interface designs.

To address these issues and better understand the inherent usability trade-offs between PGP, IBE,[1] and password-based key management, we have created implementations of each of these key management schemes. These implementations share the same overall design, differing only as needed for each different key management scheme. We are currently conducting a usability study of these prototypes using a paired-participant study methodology [7].

Related Work

Whitten and Tygar [15] conducted the first formal user study of a secure email system (i.e., PGP 5), which uncovered serious usability issues with key management and users' understanding of the underlying public key cryptography. They found that a majority of users were unable to successfully send encrypted email in the context of a hypothetical political campaign scenario. The results of the study took the security community by surprise and helped shape modern usable security research. Sheng at al. demonstrated that despite improvements made to PGP in the seven years since Whitten and Tygar's original publication, key management was still a challenge for users [13]. Furthermore, they showed that in the new version of PGP encryption and decryption had become so transparent that users were unsure if a message they received had actually been encrypted. In a separate study, Garfinkel and Miller showed that the usability of secure email could be improved by automating key management [5].

More recently, we conducted a series of user studies with Private WebMail (Pwm), a secure email prototype that tightly integrates with the Gmail web interface. In our first study, we showed that participants found Pwm to be usable, but that some users made mistakes and were hesitant to trust the system due to how transparently it secured their messages [11]. Based on feedback from this study, we modified Pwm and conducted a second user study [8]; the results of this study were highly encouraging, with Pwm rated as highly usable and no significant issues detected. We also conducted a study where we compared Pwm against three other secure email tools [7, 10]: Virtru (key-escrow), Tutanota (depot-based), and Mailvelope (PGP). Our results showed that participants preferred Pwm over the alternatives, and that even a modern PGP secure email tool is unusable. MessageGuard incorporate all the best practices learned across our studies.

In a replication of our first study of Pwm, Atwater et al. verified that participants responded positively to automatic key management [1]. They created a mock-up of Mailvelope that automatically generates keys for users, shares the public key with a key server, and automatically retrieves an email recipient's public keys as needed. Unfortunately, the mock-up did not include working key discovery, instead relying on hard-coded recipient keys. As noted in their paper, this led to a problem in the simulation of the study task, limiting the conclusiveness of the results. The PGP-based secure email tool in our study incorporates ideas proposed by Atwater et al., but is fully functional and should lead to a more accurate analysis of PGP's inherent usability.

Bai et al. recently studied variations in PGP key management [2]. Their study held the user interface constant while varying the method by which users discover PGP keys—between manual key exchange and automated key discovery through a key directory service. Their work demonstrates the effectiveness of holding user interfaces constant when evaluating aspects of key management.

[1] Identity-based encryption [12].

Data Analysis

Each system will be analyzed and compared using the following metrics:

- The System Usability Scale (SUS).
- Time taken to complete each stage of the task.
- Number of mistakes participants make. Mistakes are defined as sending sensitive information in the clear.
- Percentage of participants who prefer the system over the others tested.

Additionally, we will analyze participants' qualitative feedback in order to better understand the strengths and weaknesses of each system.

Systems

Using MessageGuard (see sidebar) we built three variants of secure email: PGP, IBE, and password-based, referred to as MG-PGP, MG-IBE, and MG-PW, respectively. The system and interface design for each system is identical except for the functionality and interfaces that must differ for each key management scheme. Each system can be downloaded separately for testing at https://messageguard.io/{pgp,ibe,passwords}.

In the remainder of this section we give a brief description of each key management scheme and the interface elements unique to that scheme.

PGP As users install MG-PGP, they are first instructed to create an account at the key server and verify ownership of their email address.[2] When a user sends a message, one of two things will occur: (1) a recipient hasn't setup MG-PGP and an email is generated that asks the recipient to install MG-PGP so that they can be sent an encrypted email; (2) all recipients have already setup MG-PGP and encryption continues without any delay.

IBE MG-IBE functions similarly to MG-PGP, except that MG-IBE can encrypt messages for recipients who have not yet installed MG-IBE because IBE public keys are just email addresses. This obviates the need for users to ask their recipients to first install MG-IBE.

Password-based Unlike MG-PGP or MG-IBE, users of MG-PW do not need to register at the key server. When these users encrypt a message, they will be prompted to generate a password that is used to encrypt the message. They are also informed that they will need to share this password with any recipients using an out-of-band method (e.g., phone call). When recipients receive a message, they must enter the password in order to decrypt it.

[2] This is necessary in order to ensure that only valid public keys are uploaded to the server.

Methodology

We are conducting an IRB-approved user study that compares the three secure email variants. Our methodology is based on a methodology we introduced in prior studies of secure email [7, 10]. This methodology tests secure email using pairs of participants to better simulate grassroots adoption. The study is a within-subjects design involving a total of 40 – 50 pairs of participants.

Study Design

When participants arrive, they provide written consent and are read a brief introduction detailing the study. Participants are informed that they will be in separate rooms during the study and will use email to communicate with each other. Participants are also informed that a study coordinator will be with them at all times, and can answer questions they might have.

During the study, participants role-play a tax preparation scenario. Participant A (hereafter referred to as Johnny) is told they need Participant B's (hereafter referred to as Jane) help with filing taxes. Johnny is also told that since he is sending sensitive information (e.g., SSN) that he should encrypt his information. Johnny is also given the URL for one of the three secure email variants we are testing. Jane is told to wait for her friend to send her an email with his tax information. Once Jane has received this information, she is instructed to (securely) reply to Johnny with a confirmation code to conclude the task.

While participants wait to receive email from each other, they are allowed to browse the Internet, use their phones, or engage in other similar activities. This is done to pro-

vide a more natural setting for the participants, as well as to avoid frustration if participants have to wait for an extended period of time while their friends figure out how to use secure email. Study coordinators are allowed to answer questions related to the study, but are not allowed to provide instructions on how to use the secure email systems. If participants become stuck and ask for help, they are invited to consider how they would solve a similar problem in the real world.

Study Questionnaire and Post-Study Interview
We administer our study questionnaire using the Qualtrics web-based survey software. At the start of the study, participants answer a set of demographic questions. Afterwards, participants complete the study task for each of the three systems. The order in which they use each system is randomized.

Upon completing the study task for each system, participants are asked several questions related to their experience with that system. First, participants complete the ten System Usability Scale questions [3, 4]. Second, participants are asked to describe in their own words what they liked about the secure email system, what they would change, and why they would change it.

After the completion of the survey, study coordinators conduct a post-study interview with their respective participants. Study coordinators focus on issues that had arisen during the study and probe for more details regarding areas of confusion. Coordinators also assess how well users understand the very different security models inherent in the different key management schemes.

After the participants complete their individual post-study interviews, they are brought together for an additional post-study interview. In this group interview, participants are asked to discuss their experience with each other. Additionally, they are asked to describe how their ideal secure email system would function; while participants are not system designers, our experience has shown that when asked to design ideal systems, participants often reveal preferences that otherwise remain unspoken [11, 7].

Study Setup
We are gathering participants using Craigslist and by distributing posters broadly across our local university. Participants are required to be accompanied by a friend, who serves as their counterpart for the study. To standardize participants' experience, both participants are required to have and use their personal Gmail account.

Participants are given sixty minutes to complete the study, with about 35-40 minutes allocated for task completion. Both participants are compensated $15 USD for their participation.

Study Limitations

Our study has limitations common to all existing secure email studies. First, our populations are not representative of all groups, and future research could broaden the population (e.g., outside the USA, non-Gmail users). Second, our study was short-term, and future research should look at these issues in a longer-term, longitudinal study. Third, our study is a lab study and has limitations common to all studies run in a trusted environment [6, 14].

In our study we only examine the case where a single user sends email to one other user. While this is the most likely scenario for secure email among the masses, future work could also explore alternative scenarios (e.g., sending to multiple users, mailing lists).

Conclusion

It is unclear how the inherent usability of various key management schemes compare against each other. To better understand this question, we have developed three secure email tools that use PGP, IBE, and password-based encryption, respectively. Each of these systems are developed using MessageGuard, ensuring that they collectively have a consistent system and interface design, differing only in ways that are intrinsic to each key management philosophy. This is the first study to directly compare wholly different key management schemes while holding the interface largely constant. We believe that this will provide us with valuable insights on the inherent usability trade-offs between each approach.

REFERENCES

1. Erinn Atwater, Cecylia Bocovich, Urs Hengartner, Ed Lank, and Ian Goldberg. 2015. Leading Johnny to Water: Designing for Usability and Trust. In *Eleventh Symposium On Usable Privacy and Security*. 69–88.
2. Wei Bai, Moses Namara, Yichen Qian, Patrick Gage Kelley, Michelle L. Mazurek, and Doowon Kim. 2016. An Inconvenient Trust: User Attitudes Toward Security and Usability Tradeoffs for Key-Directory Encryption Systems. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*.
3. John Brooke. 1996. SUS — a quick and dirty usability scale. In *Usability Evaluation in Industry*. CRC Press.
4. John Brooke. 2013. SUS: A Retrospective. *Journal of Usability Studies* 8, 2 (2013), 29–40.
5. Simson L Garfinkel and Robert C Miller. 2005. Johnny 2: A User Test of Key Continuity Management with S/MIME and Outlook Express. In *Proceedings of the First Symposium on Usable Privacy and Security*. ACM, 13–24.
6. Stanley Milgram and Ernest Van den Haag. 1978. Obedience to authority. (1978).
7. Scott Ruoti, Jeff Andersen, Scott Heidbrink, Mark O'Neill, Elham Vaziripour, Justin Wu, Daniel Zappala, and Kent Seamons. 2016. "We're on the Same Page": A Usability Study of Secure Email Using Pairs of Novice Users. In *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems*. ACM.
8. Scott Ruoti, Jeff Andersen, Travis Hendershot, Daniel Zappala, and Kent Seamons. 2015. Private Webmail 2.0: Simple and Easy-to-Use Secure Email. *arXiv preprint arXiv:1510.08435* (2015).
9. Scott Ruoti, Jeff Andersen, Tyler Monson, Daniel Zappala, and Kent Seamons. 2016. MessageGuard: A Browser-based Platform for Usable, Content-Based Encryption Research. *arXiv preprint arXiv:1510.08943* (2016).
10. Scott Ruoti, Jeff Andersen, Daniel Zappala, and Kent Seamons. 2015. Why Johnny Still, Still Can't Encrypt: Evaluating the Usability of a Modern PGP Client. *arXiv preprint arXiv:1510.08555* (2015).
11. Scott Ruoti, Nathan Kim, Ben Burgon, Timothy van der Horst, and Kent Seamons. 2013. Confused Johnny: when automatic encryption leads to confusion and mistakes. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*. ACM.
12. Adi Shamir. 1985. Identity-based cryptosystems and signature schemes. In *Advances in cryptology*. Springer, 47–53.
13. S. Sheng, L. Broderick, CA Koranda, and JJ Hyland. 2006. Why Johnny Still Can't Encrypt: Evaluating the Usability of Email Encryption Software. In *Proceedings of the Second Symposium On Usable Privacy and Security – Poster Session*.
14. Andreas Sotirakopoulos, Kirstie Hawkey, and Konstantin Beznosov. 2010. "I did it because i trusted you": Challenges with the study environment biasing participant behaviours. In *SOUPS Usable Security Experiment Reports (USER) Workshop*.
15. Alma Whitten and J. D. Tygar. 1999. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. In *8th USENIX Security Symposium*. `citeseer.nj.nec.com/whitten99why.html`