Secure Email - A Usability Study

Projet Fin d'Études (PFE)
CryptogrAphie, Sécurité, et vie Privée dans les Applications et Réseaux (CASPAR)
Supervisor: Karima Boudaoud
Polytech Nice Sophia Antipolis

Adrian Reuter
Technische Universität München (TUM)
Email: adrian.reuter@tum.de

Ahmed Abdelmaksoud
Polytech Nice Sophia Antipolis
Email: ahmed-adel.abdelmaksoud@etu.univ-cotedazur.fr

Wadie Lemrazzeq
Polytech Nice Sophia Antipolis
Email: wadie.lemrazzeq@etu.univ-cotedazur.fr

ABSTRACT

Keywords

Source Routing, Source Packet Routing, Segment Routing, MPLS, SPRING, RPL, DSR, RH0

1. INTRODUCTION

The most valuable information that is collected and traded in today's internet, is the personalized data of internet users. To prevent cyber-crime and to protect user privacy, almost all services running in the internet critically depend cyber security measures. Most dominantly, transport layer security (TLS) is used for securing a variety of communication protocols. Particularly when browsing the world wide web using HTTP, transport security has found wide adoption and awareness of its imperative necessity. Internet Banking, shopping on Amazon.com, accessing governmental e-services - those are just a few examples for internet use cases in which users became more and more aware of the security critical nature of these web applications, even if they might not understand the attack vectors and their risks in depth.

A huge step towards more secure internet communication has been the integration of end-to-end cryptography in mobile internet messenger services such as Whatsapp, Signal or Telegram. In contrast, for securing one of the most commonly used communication channels - the electronic mail - end-to-end encryption is only applied by a neglectable faction [1] of email users. Standardized technologies for cryptographically securing email exchanges have been available for decades. Nevertheless most users rely on unencrypted and unauthenticated email communication, often without being aware that there exist mechanisms which would mitigate the security implications that come with it. The relevance of applying end-to-end cryptography to our daily email communication can be exemplarily depicted with the following scenarios:

- Protecting confidentiality
  The content of an email should be kept confidential and should not be readable by someone other than the intended recipient of the email. Emails often communicate sensitive data about human or non-human assets. For example: Personal data, business secrets, industrial know-how, investigative journalism and almost all password-recovery routines for any web application critically depend on email exchanges.
- Protecting privacy
  Personal information of users communicated over emails should not be processed and analyzed by any other entity than the intended recipient of the email. For example: Entities that can listen to internet traffic such as internet service providers, email providers themselves, analytics services should not be able to collect personal data exchanged via email, as the content was not intended to be exposed to 3rd party entities
- Protecting integrity
  The content of an email should not be tampered with by any other entity other than the sender of an email, i.e. no other entity should be able to altering the content without the recipient being able to detect this illegitimate modification. This feature is typically provided by a digital signature.
- Provide authenticity and non-repudiation No entity should be able to impersonate another email user and write emails in his or her name. In order words, the recipient of an email should be able to trust the origin of an email; a received email should not originate from any other entity than the sender indicated to the recipient. Likewise, a sender should not be able to disclaim the content of sent mails, which provides non-repudiation for the recipient. For example: Information distributed via mail (e.g. agreements on deliverables, meeting appointments, conditions and contract sent within email as document in email attachments) cannot neither be changed nor denied afterwards. Emails are authentic and are not spoofed, severely complicating phishing attacks.

To achieve the previously mentioned security goals, two major end-to-end encryption technologies exist since decades,

namely Pretty Good Privacy (PGP) [2] and Secure Multipurpose Internet Mail Extensions (S/MIME) [3]. A recent initiative called Pretty Easy Privacy (pEp) [4] made efforts to simplify the usage of end-to-end cryptography in email communication for novice users. Unfortunately, those technologies are still barely deployed. According to [1] more than 95% of the overall email traffic is exchanged without end-to-end encryption.

Our central research questions are:

- identifying why users are hesitating to use the above mentioned technologies.
- which usability issues exist that hinder users from securing their daily email communication using end-to-end encryption.

2. RELATED WORK

@Wadie

3. ANALYSIS OF END-TO-END ENCRYPTION TECHNOLOGIES FOR EMAILS

3.1 Pretty Good Privacy (PGP)

Developed by Phil Zimmermann in 1991, Pretty Good Privacy (PGP) is an encryption program that provides cryptographic privacy and authentication for data communication. PGP is used for signing, encrypting, and decrypting texts, e-mails, files, directories, and whole disk partitions and to increase the security of e-mail communications. In this study we will focus only on using PGP for e-mail security (e-mail encryption). It follows the OpenPGP standard for encrypting and decrypting data. Many e-mail clients provide OpenPGP-compliant e-mail security as described in RFC 3156 [5]. The current specification is RFC 4880 (November 2007) [6]. PGP encryption uses a serial combination of hashing algorithms (SHA-1, SHA-224 / 256 / 384 / 512), data compression algorithms (zip, zlib, and bzip2), symmetric encryption algorithms (3DES, AES-128 / 192 / 256, CAST5, IDEA) and finally asymmetric encryption algorithms (ElGamal, RSA (MUST NOT <1024 bits)).

In PGP, one-off key is generated randomly, which is known as the session key. The session key encrypts the message, which is the bulk of the data that needs to be sent. This type of encryption is relatively efficient, but it has a problem of sharing the session key with your recipient because without the key your recipient will only see the ciphertext. PGP solves this problem with public-key cryptography, also known as asymmetric cryptography. In this kind of encryption there are two keys: a public key and a private one. The public key of your potential correspondent can be found by searching through key servers or by asking the person directly. Moreover, each public key is bound to an e-mail address and has a unique fingerprint which can be used to get the right corresponding public key [7]. In PGP, public-key encryption isn't used to encrypt the message, just the one-off session key that was generated to encrypt it. It would take too long and use a larger amount of computational resources. Since the body of the message usually contains the bulk of the data, PGP uses the more economical symmetric-key encryption for this. It reserves the lumbering public-key encryption for the session key, making the whole process more efficient. Our written signatures are frequently used to verify that we are who we say we are. They are far from foolproof, but they are still a useful way of preventing fraud.
Digital signatures are similar, using public-key cryptography to authenticate that the data comes from the source it claims to and that it has not been tampered with. Digital signatures work by using an algorithm to combine the sender's private key with the data that they are authenticating. The plaintext of the message is fed through a hash function, which is an algorithm that transforms inputs into a fixed-size block of data, called a message digest. The message digest is then encrypted with the sender's private key. This encrypted message digestis what is known as the digital signature [8]. In PGP, the digital signature is sent alongside the message body (which can either be encrypted or in plaintext). When someone receives a digitally signed electronic mail (e-mail), they can check its authenticity and integrity by using the public key of the sender. First, a hash function is used on the message that was received and this gives the message digest of the email in its current form [8]. The next step is to calculate the original message digest from the digital signature that was sent. The sender's public key is used to decrypt the digital signature, and this gives the message digest exactly as it was when it was signed by the sender. The final step is to compare the message digest from the email they received to the message digest that they derived from the digital signature. If the message has been altered, then the message digests will be completely different, and the recipient will know that there is a problem with the message. If the two message digests are not identical, there are three likely culprits [8]:

- The public key used to decrypt the digital signature was not linked to the private key that was used to encrypt it. This means that the sender may not be who they say they are.
- The digital signature may be fake.
- The message has been changed since it was signed.

In addition to all these security measures, another important factor is the emails attachments. PGP can also be used to encrypt your attachments. There are a two of ways to do this:

- One approach to encrypt an email with PGP is to encrypt everything separately. This means that the message body and attachments are individually encrypted and signed. ItâĂŹs called PGP/Inline [9].
- A newer approach is PGP/MIME, which - in contrast to PGP/Inline - encrypts and signs the message as a whole, including attachments [9].

Finally, PGP relies on the concept of web of trust to establish trust between the users. it is critical that the public key used to send messages to someone or some entity actually does 'belong' to the intended recipient [5]. Simply downloading a public key from somewhere is not a reliable assurance of that association. The web of trust grew as a

way of vetting that each PGP public key and user are really connected to the person or organization that they are said to represent. The web of a trust connects the real-life entity with their public key by using a third party to sign the userâĂŹs PGP public-key and it does it all without a central authority that can collapse or be corrupted [6]. If a user knows another PGP user personally, he can confirm that their public key is linked to their actual identity. He can put his trust in them and digitally sign his public-key, which shows that at least one person vouches for his identity. Then, he can also do the same. Over time, this builds an interconnected web of trust, with lots of people vouching for each other with digital signatures that verify their ownership of a public key.

In conclusion, PGP encryption leverages a range of techniques to provide secure and private email communications. These include compression, public-key encryption, symmetric encryption, digital signatures, and the web of trust. Together, these allow its users to send encrypted messages in an efficient manner. It also let them check whether a message is authentic and has not been altered. The OpenPGP standard was formed so that everyone could use it for free, helping to make it the most widespread form of email encryption and protect the privacy of email users.

3.2 Secure Multipurpose Internet Mail Extension (S/MIME)

With the increase of power in terms of computation and networking capability, peopleâĂŹs need for using non-textual objects such as image, audio, and video in emails became a fair demand. To support diversity of content, multipart message structure, and non-English text, the Multipurpose Internet Mail Extension (MIME) was proposed in 1993 [8]. S/MIME (Secure/Multipurpose Internet Mail Extension) is an enhancement of MIME to provide cryptographic security for the MIME based emails [8]. To better understand S/MIME, let us first take look at MIME. S/MIME adds some additional content types to the MIME to provide security services. The current S/MIME version 3.2 obsoletes all earlier versions. However, most implementations still bear version 3.1 features for digital signature processing. Today, popular email clients such as MS Outlook 2013, MS Outlook 2016, Mozilla Thunderbird, Apple Mail support S/MIME enabled messages. S/MIME version 3.2 is defined by the following five major specifications.

- Cryptographic Message Syntax (RFC 4853)
- Cryptographic Message Syntax (CMS) Algorithms (RFC 3370)
- Diffie-Hellman Key Agreement Method (RFC 2631)
- S/MIME Version 3.2 Certificate Handling (RFC 5750) [10]
- S/MIME Version 3.2 Message Specification (RFC 5751) [11]
- Enhanced Security Services for S/MIME (RFC 2634) [12]

Cryptographic Message Syntax (CMS) To define how security services, such as confidentiality or integrity, can be added to MIME content types, S/MIME defines Cryptographic Message Syntax (CMS). The syntax in each case defines the exact encoding scheme for each content type. Discussed below are the different types of messages and different sub-types that are created from these messages [14].

- Data Content Type is an arbitrary string. The object created is called ”Data”.
- Digested-Data Content Type is used to provide integrity for the message. The result is typically used as the content for the enveloped-data content type. The encoded result is an object called ”digested-data”. The process of creating ”digested-data” involves the following two steps
  1. Using the hash algorithm of the user's choice a message digest is created from the content.
  2. The message digest, the algorithm, and the content are added together to create the ”digested-data” object.
- Signed Data Content Type provides authenticity and integrity of data. It contains any type and zero or more signature values. The encoded result is an object called ”signed-data”. A signed data message can only be viewed by the recipient of with S/MIME capability. The following are the steps in the process.
  1. For each signer, a message digest is created from the content using the specific hash algorithm chosen by the signer.
  2. Each message digest is signed by the signer with his or her private key.
  3. The content, signature values, certificates, and the algorithms are then collected to create the ”signed-data”.
- Enveloped-Data Content Type is used to provide privacy for the message. It contains encrypted content of any type, and zero or more encrypted keys and certificates. The encoded result is an object called ”enveloped-data”. The steps involved in the process is are as follows
  1. A pseudo-random session key is created for the symmetric-key algorithm to be used to encrypt the content.
  2. The content is encrypted using the defined algorithm and created session key.
  3. For each recipient, a copy of the session key is encrypted with the public key of the recipient.
  4. The encrypted contents, encrypted session keys, algorithm used, and certificates are encoded using Radix-64.

Cryptographic Algorithms S/MIME specifies several cryptographic algorithms for use. The term ”must” indicate absolute requirement, and the term ”should” imply recommendation only. S/MIME recommends using SHA-256 as hash

function for creating message digests. For content encryption RSA is recommended.

Key Management and Certification The key management in S/MIME is a combination of key management used by X.509 and PGP. S/MIME uses public key certificates signed by the Certification Authorities (CA) defined by X.509. The user is responsible for maintaining the web of trust to verify signatures as defined by PGP. In general, for key establishment, a user needs to get his or her pair of public and private keys certified by trusted CAs [13].

3.3 Pretty Easy Privacy (pEp)

4. METHODOLOGY

In this chapter we present our approach for evaluating the usability of the three end-to-end encryption technologies examined in this project. We decided to follow a threefold approach:

1. We identify the most commonly Mail User Agents (MUA) also known as email programs, that - natively or by additional plugins - support at least one of the three technologies PGP, S/MIME or pEp. We assess the usability of the encryption features in each of these mail programs **ourselves**, to get a personal impression as well as to anticipate the challenges that other users might face when cryptographically securing their emails.

2. We prepare, execute and evaluate an **online survey**, which assesses the usability of PGP, S/MIME and pEp of a broad audience.

3. We conduct **live observation tests**, in which we let participants use PGP, S/MIME or pEp to write cryptographically securing emails. We observe the participants journey of installing, configuring and using the encryption features up to the point where we receive an email that was successfully encrypted and signed. We do these live observations to:

    - validate the responses to the questions of our online survey.
    - get a more reliable and precise feedback of the participants, revealing exactly which configuration step constitutes a challenge or which aspects hinder participants to apply email encryption in their regular mail exchange.

For the assessment of the three technologies, we divided it as following: one is responsible of reading the white-paper and two are responsible of configuring and testing the technology on every platform available. In same way for the live observations, we had twelve tests, each one of us took a number of tests and he was responsible to find a user to do it as well as supervise it.

4.1 Online survey

Goal of the online survey The aim of this online survey on email end-to-end encryption technologies was threefold. First, to explore users understanding and awareness of security in emails exchanges, their expectations and opinions on end-to-end encryption. Second, to establish a pattern about the propagation of the three technologies existing in the market. third, to compare this online survey which is quantitative data with the live survey defined as a qualitative data then validate those two surveys.

Structure Web based survey was our choice in order to collect data, because of its advantage, low cost and quick distribution. For the survey design, we tried to stimulate the participant to be objective, by keeping control on the flow of questions, for example, before asking specific question on each technology, we ask the participant if they have any knowledge of it. We tried to keep also an unbiased flow, by avoiding framing bias, through proper wording, and the use of clear, unambiguous and concise wording, in order to let the participant only depends on his personal knowledge on the topic. The survey included closed-ended questions (multiple choice questions), open-ended questions and ranked questions with a balanced rating scale. The online survey treats six sections, which one has its own purpose. Also, we included section skipping logic based on the participant answers.

- The first section introduces the survey for the participant and handles the demographic data, but first it introduces the participant to the survey by explaining the aim of the survey, thenceforth it asks regular demographic questions and the type of the work organization to conclude if there is any relation between the need of secure email communications and the activity work type.

- The second section interact with the participant experience on the field of email exchange, this section helps us basically to identify the spread of email encryption through the participants email exchange, also make a link between the spread of email encryption and email clients usage.

- The third section evaluate the participant experience with PGP, basically if the participant had already an experience with it, we discuss with him the advantage and the disadvantage of the technology related to its implementation on the email client based on our study on email client supporting end-to-end encryption.

- The fourth section evaluate the participant experience with S/MIME, mostly the same questions as the previous section adjusted for the technology. The fifth section evaluate the participant experience with pep, this section is shorter than the previous ones by reason of being new to the market, it includes some of the previous questions also adjusted to the technology.

- The last and sixth section, gather the overall impression on end-to-end encryption, by scaling the degree of awareness of the participants on matter of the security of email exchange especially if they had an experience with an email piracy issue.

By the end we would like to mention that, the majority of those questions, specifically those related to technologies were raised when we conduct our study with the email clients which implement end-to-end encryption.

4.2 Live observations

5. RESULTS

5.1 Online Survey

We note that the given time and resource constraints limited our sample size, and consequently the number of survey participants is too low to permit informative statistical analysis of the results. However, we do not consider this to be a major shortcoming since demonstrating statistical significance is not essential for the purpose of this study. The number of participants was sufficient for our purposes.

The online survey was launched on 30 November 2018, it reached 50 participants on 12 December 2018 when we start analysis of surveys results.

As we described before, the online survey has six section, each section treats a certain matter related to end-to-end encryption usability. The survey begins with a demographic section. We can summarize that the majority of the participant was under 30 year old, coming from Germany, Egypt, Morocco, most of them were students, and employees working for IT organizations.

Concerning their personal experience with email exchange, it appears that emails constitute a remarkable portion of their daily communications, reaching at least 7 emails per day, but most of them are nor encrypted neither signed, 38% receive at least 1 email encrypted per day (Figure 3), and less than half of the our survey participants were obliged to use end-to-end encryption by their organizations.

The participants use different platforms and MUAs, more than half use dedicated mobile applications, 50% use webmail, and 44% use dedicated desktop applications (Figure 4). From our participants, 40% state they knew PGP prior to this survey, 16% are also using it (Figure 5).

70% of our participants state that they cannot use PGP on all mails due to the fact that the recipient does not use PGP. On the other hand, 25% thinks that itâĂŹs difficult to find the recipient's public key, 20% think configuring PGP is time consuming and just 5% declare that PGP it is not implemented on their favorite platform / email client. 20% of our participants always verify the fingerprint of the recipient key, 30% do so occasionally and 35% never do (Figure 6). The participants concede that PGP guaranties privacy, confidentiality, authenticity and integrity, adding that it has no cost in order to use it. On the other hand they disprove as being complex, comparing fingerprints was difficult and time consuming, and requiring the recipient to use it as well, which is not always the case given that PGP is not widely adopted. Also, participants suggest to make PGP supported on all platforms and simplify fingerprint comparison.

Also for S/MIME, we asked before proceeding with detailed questions, if our participants already knew S/MIME, with 36% saying so (Figure 7). The participants experience also that the recipient does not use S/MIME with 61%, 28% do not trust digital certificates or its issuing entity, and only 11% do not know how to obtain digital certificate. 17% encounter difficulties configuring their environment to use S/MIME. 27% admit that they had issues with untrusted certificates (Figure 8), for 28% of the participants the fact that they have to pay for a trustworthy certificate is an obstacle. The participants agreed that S/MIME has the advantage of being integrated in most email clients including Apple MacOS/iOS, but they discredit S/MIME because they need to pay to obtain a trustfully certificate.

Apparently, pEp it is not as known as the other previous technologies, only 10% who know it. No participant stated that he ever used it (Figure 9). 40% of the participants hesitate to use pEp because their recipients will not use pEp. Assessing their overall impression, the participants are mostly aware of the importance of email encryption: 66% think that email encryption is important to very important (Figure 10). Considering the scenario of non-secured mail exchange, more than 60% of our participants can imagine that their emails can be passively or actively tampered with; an even larger percentage of 86% assumes that an entity other than the mail recipient can read mail contents (Figure 11). Assessing the importance of specific security goals, almost all of our participants estimate the confidentiality, integrity and authenticity of their mails as important or very important, only for 6% confidentiality does not matter and only for 2% the integrity of their sent mails does not matter (Figure 12).

5.2 Live observations

6. DISCUSSION

7. CONCLUSION

With our project, we identified the most frequent usability issues that users face when protecting their email communication using Pretty Good Privacy (PGP), Secure Multipurpose Internet Mail Extension (S/MIME) or Pretty Easy Privacy (pEp). First, we did a technical analysis of those three technologies, elaborating on trust establishment, key exchange, cryptographic algorithms and provided security features. Secondly we prepared and launched an online survey, aiming for a broad audience to collect a maximum feedback. The survey does not only identify the usability issues of each technology, but also assesses the general impression of our audience towards the importance of email encryption. Third, we conducted live observations in which participants were to install, configure and use either PGP, S/MIME or pEp in presence of a team member that observed the participant testing the technologies and that provided support in case the participant encountered difficulties. Our three fold approach gave us both, an overall view on the awareness of the email users as well as a detailed view on the causes for which users are hesitating to use the mentioned technologies. The overall impression we received from the online survey showed us that the email users are aware of the importance of email encryption with 32% saying it's very important. Additionally, users are very concerned about identity theft, as 78% of the participants want to make sure that no other person is able to write email using their name and 80% of the participants want to be sure that the content of their mail isn't changed by someone else while being transferred to the recipient. It shows that for many users, signing emails is more important than encrypting them. For the usability issues we found, we propose some sample improvements that we suppose will make it easier for the users to apply PGP, S/MIME or pEp in their email communication.

We present the improvement suggestions by technology:

- Improvements for PGP:
  We found out that on all tested platforms, users are restricted to the search for the recipient key only at one key server at a time, which makes importing the recipient key the major obstacle when using PGP. Thus, we suggest to fix this issue by letting implementations search for the recipient key on all available key-servers

at a time, without the user needing to manually adjust the key server for key import

- Improvements for S/MIME
  We suggest Certification Authorities to send further information on how to import the certificate into the most frequently used email programs, alongside the email which contains the requested digital certificate. Also, we suggest to integrate S/MIME support into webmail plugins, as webmail nowadays is used more commonly than dedicated desktop applications.

- Improvements for pEp
  We suggest pEp implementations to add further explanation on what to do with the trustwords displayed to the users during handshake. We furthermore advise to briefly explain the pEp color scheme that represents the security status of a communication channel to the user. Moreover we criticize the design choice of not being able to do a new handshake with a user whose public key was previously mistrusted. As a trustword mismatch can too easily occur due to a different language selection of the trustword dictionary on side of both users.

8. REFERENCES

[1] Matthew Green, "The Daunting Challenge of Secure E-mail." `https://www.newyorker.com/tech/annals-of-technology/the-daunting-challenge-of-secure-e-mail`; last accessed on 2019/02/13.

[2] D. Atkins, W. Stallings, and P. Zimmermann, "PGP Message Exchange Formats." RFC 1991 (Informational), Aug. 1996. Obsoleted by RFC 4880.

[3] B. Ramsdell, "S/MIME Version 3 Message Specification." RFC 2633 (Proposed Standard), June 1999. Obsoleted by RFC 3851.

[4] pEp Security, "Pretty Easy Privacy." `https://www.pep.security/`; last accessed on 2019/02/13.

[5] M. Elkins, D. D. Torto, R. Levien, and T. Roessler, "MIME Security with OpenPGP." RFC 3156 (Proposed Standard), Aug. 2001.

[6] J. Callas, L. Donnerhacke, H. Finney, D. Shaw, and R. Thayer, "OpenPGP Message Format." RFC 4880 (Proposed Standard), Nov. 2007. Updated by RFC 5581.

[7] Wikipedia, "Pretty Good Privacy." `https://en.wikipedia.org/wiki/Pretty_Good_Privacy/`.

[8] Josh Lake, "What is PGP encryption and how does it work?." `https://en.wikipedia.org/wiki/Pretty_Good_Privacy/`; last accessed on 2019/02/13.

[9] Protonmail, "What are PGP/MIME and PGP/Inline?." `https://protonmail.com/support/knowledge-base/pgp-mime-pgp-inline/`; last accessed on 2019/02/13.