Systematization of Threats and Requirements for Private Messaging with Untrusted Servers: The Case of e-Mailing and Instant Messaging

Iraklis Symeonidis[a] and Gabriele Lenzini[b]
SnT, University of Luxembourg, Luxembourg
{*firstname.lastname*}*@uni.lu*

Keywords: Security, Privacy, Threat Modeling, Decentralization, System Model, Private Messaging.

Abstract: Modern email and instant messaging applications often offer private communications. In doing so, they share common concerns about how security and privacy can be compromised, how they should face similar threats, and how to comply with comparable system requirements. Assuming a scenario where servers may not be trusted, we review and analyze a list of threats specifically against message delivering, archiving, and contact synchronization. We also describe a list of requirements intended for whom undertakes the task of implementing secure and private messaging. The cryptographic solutions available to mitigate the threats and to comply with the requirements may differ, as the two applications are built on different assumptions and technologies.

1 INTRODUCTION

Long stating facts have made clear to the public that governments routinely collect and invest in the ability to massively eavesdrop private messages of citizens (Stanger, 2019). In response to such revelations, and sensitized to preserve the "right to privacy" (*e.g.*, see (European Parliament and the European Council, 2016)) beyond trusting authorities, messaging servers, and communication channels, an increasing number of companies are investing in *private messaging* solutions (Unger et al., 2015). Private messaging can be realized, for instance, in E-mailing applications (*e.g.*, *Protonmail*, $p{\equiv}p$ (Marques et al., 2019)). It can be offered in instance messaging (*e.g.*, *Signal*, *Telegram*).

Whatever the hosting system, E-mailing or a Instant Messaging (IM), private messaging should ensure that, in an written exchange between peers, no one but the sender and the intended receiver(s) are capable of reading the messages at any time: past, current, and future.

Realizing this goal can be challenging despite the many off-the-shelves security APIs and libraries today available for software engineers. In the two types of systems, private messaging may realize different mitigation mechanisms or satisfy different design choices. For instance, confidentiality can be preserved in multiple ways and with various cryptographic primitives and, in a situation where lack of privacy may endanger one's life, such as in investigative journalism, such solutions are expected to be more stringent than those called for in situations where users relaxedly exchange messages with a broad public.

It must be said that the urgency of realizing private messaging is lively debated. As it has been already established for paper mail, *secrecy of correspondence* is regarded as a fundamental legal principle of democracies in many countries (Marotta and Russell, 2013). It is often argued that preserving the privacy of communication between humans should be taken for granted; that it should be respected and protected as a fundamental human right (Council of Europe: European Court of Human Rights, 2016); and that the communication of peers has to be protected in a self-evident manner. An ongoing discussion, mainly following the revelation by Edward Snowden and other whistleblowers (Stanger, 2019), evaluates whether protections should be delivered by the offering technology and independent of what the user wants to write or why. In this discussion, E-mailing and messaging systems are regarded as the central and most prominent digital communication means using which to realize that legal principle.

The problem is that composing security components without a clear understanding of the threats leads to unclear or incomplete privacy guarantees. Without that understanding, an application is likely

[a] https://orcid.org/0000-0003-4494-6685
[b] https://orcid.org/0000-0001-8229-3270

to need to be patched to fix vulnerabilities discovered in the meantime (Dkg, 2019; Rijhansen, 2019), and this is not optimal. Better, it is an approach capable of ensuring *privacy-by-design*: here the appropriate workflow demands first deciding against what threats one intends to uphold and then listing the related functional requirements, to fulfil which appropriate solutions are selected.

Contribution. In reference to both E-mailing and IM, we describe the entities involved, the operations, and the system features of such applications. While discussing similarities in the two types of applications, we highlight the differences. Considering users and systems as assets and assuming that servers may be untrustworthy, we also orderly catalogue threats against privacy messaging, and discuss security requirements to mitigate them.

As such, this work offers a systematization of threats and requirements for use of experts. The work aim to assist them to assess current E-mailing and IM applications and to design implement more robust private messaging systems. Although limited to three specific protocols in private messaging (*messages exchange*, *search and archiving*, and *contact synchronization*), our list of threats and requirements carry already a quite rich set of challenges for whomever aims to realize private messaging in E-mailing and IM.

It is also an introduction to private messaging for non-experts.

2 BACKGROUND AND RELATED WORK

E-mailing and IM are two types of messaging applications. E-mailing depends on Simple Mail Transfer Protocol (SMTP) for email transmission (Klensin, 2008). SMTP has been designed with no built-in security. As a result, there is no prevention from adversaries aiming for eavesdropping, spoofing, and impersonating email messages and addresses. IM instead, depends on Extensible Messaging and Presence Protocol (XMPP), a near-real-time messaging protocol consisting of relatively small chunks of Extensible Markup Language (XML) structured data between network endpoints (Saint-Andre, 2004). For XMPP, Transport Layer Security (TLS) offers channel encryption between the application peers, for instance a user client and its IM server.

E-mailing stands for high-latency and asynchronous communication in message delivery. It means that the communicating parties are not required to be simultaneously online (Resnick, 2008). IM is traditionally considered as low-latency and synchronous, with the peers required to be online in order to exchange messages (Saint-Andre, 2011). However, nowadays certain instant messaging applications allow high-latency and asynchronous message delivery. WhatsApp and Signal (Signal, 2013) as examples of new generation IM. Despite the differences, messaging applications offer common features such as transmitting messages, archiving, searching, and optionally recipient's identity verification and contacts list management.

For E-mailing and IM, there are protocols that enable users to exchange messages in a secure and privacy-enhancing way. Clark et al. (Clark et al., 2018) lists the state of the art protocols for secure E-mailing and Shirazi et al., (Shirazi et al., 2018) for privacy with the main focus on anonymous remailers. Unger et al. (Unger et al., 2015) list protocols for secure messaging, and Ermoshina et al. (Ermoshina et al., 2016) provides a short overview of decentralized E-mailing and IM applications. To the best of our knowledge, there is no existing work that methodologically analyzes the security and privacy threats and corresponding requirements for private messaging.

3 METHODOLOGY

For a system model analysis in private messaging we extracted the entities involved and the features that a system considers by following the work of (Saint-Andre, 2011; Klensin, 2008; Unger et al., 2015; Clark et al., 2018; Ermoshina et al., 2016; Shirazi et al., 2018) (see Fig. 1).

With a focus on private messaging considering untrusted servers, we compile the list of security and privacy threats for both E-mailing and IM. To distinguish and group the classes of threats and requirements we apply the STRIDE threat modeling (Howard and Lipner, 2009; Microsoft, 2010) for security and LINDDUN (Deng et al., 2011) for privacy. Both frameworks are used by the industry and the research community nowadays (Danezis et al., 2015; Symeonidis et al., 2017).

We need to stress that we extended LINDDUN threat modeling in two different ways. We incorporate *content unawareness* to *policy and consent non-compliance* threat. Note that, *content unawareness* is considered as a principle of *policy and consent non-compliance* under General Data Protection Regulation (GDPR) (European Parliament and the European Council, 2016), and thus we consider under the

policy requirement. Moreover, we extended LINDDUN to include the *privacy interdependence* privacy threat. There are scenarios where the privacy of individuals is bound to be affected by the decisions of others (Biczók and Chia, 2013). For example, a recipient of an email can forward a message to others without the original sender's notification and consent. Moreover, in messaging systems such as on WhatsApp, users can upload their contacts list in messaging servers. It enables the servers to learn and extract information about the network of users, enabling privacy interdependence.

4 SYSTEM MODEL

This section outlines a private messaging system. It describes the entities involved, the information as user assets that can be collected by the system, and the features of a private messaging system. We follow the requirements extracted from real world systems and applications as well as from the academic literature for E-mailing and instant messaging (Unger et al., 2015; Clark et al., 2018; Ermoshina et al., 2016).

4.1 Operations and Entities

We describe a set of operations and corresponding entities that are involved in E-mailing and IM messaging systems.

Users as Sender and Receiver(s): The communicating parties who exchange messages, typically referred to as senders and receivers.

Trust Establishment (TE): Supports Operations for: (i) authenticating a user, (ii) storing and distributing cryptographic material as well as (ii) synchronizing contacts of users. It utilizes *identity management*, *key management* and *contact management* servers. Note that, the operations that occur for TE are system dependent and can vary.

Message Exchange (ME): Supports operations for message delivery and synchronization. It utilizes E-mailing and IM servers for the delivery and synchronization of messages between users *i.e.*, the sender and receiver(s).

Networking Nodes (NetND): is referred to all the Internet nodes for interconnected users and E-mailing/IM servers.

Third Party: is any other entity interacting with the messaging system.

4.2 System Features and User Assets

In addition to enabling communication, *i.e.*, sending and receiving messages and attachments, there is a minimum set of common features that both E-mailing and IM systems should offer: (i) *Archive and Search* through messages and attachments; (ii) *Contacts* synchronization, and search; (iii) *Group Messages* exchange; (iv) *Multi-device* synchronization of messages and contacts across devices. In this work, we focus on threats and requirements for messages exchange, archiving/search and contact synchronization.

Considering untrusted servers, the following assets of users needs to be protected in private messaging systems: (i) *Identity and Identification Material* of sender/receiver such as e-mail, phone numbers, and public keys; (ii) *Contacts List* consisting of but not limited to e-mail, phone numbers, names, and public keys; (iii) *Content* as text and attachments; (iv) *Metadata* consisting of but not limited to the identity of the sender/receiver, packet size, and timing.

5 THREATS AND REQUIREMENTS

5.1 Adversarial Model

An adversary is any entity that leverages threats against a system. The goal of an adversary is to gain improper access to messages or any information from or about the users. An adversary can be anyone who is involved in communication such as, the users, the TE and ME entities (*e.g.*, the E-mailing and IM servers), the Networking Nodes (NetND), the Third Party (TP), or even a subset of all those entities.

Types of an adversary can be distinguished depending on whether the entity is a passive eavesdropper or an active participant in the communication and on whether it is internal to the private messaging system or an external TP one (see Table 1).

Table 1: **Adversarial Model**: passive/active and internal/external adversaries on user/system assets.

	Passive	**Active**	
	User assets		**System assets**
Internal (TE, ME, User)	read (server)	read + write (server)	execute (server)
External (TP)	read (link)	read + write (link)	availability attack (server)

Passive *vs* Active. A *passive* attacker can only eavesdrop (*i.e.*, read) messages and off-line process them. It cannot interfere with the execution of the

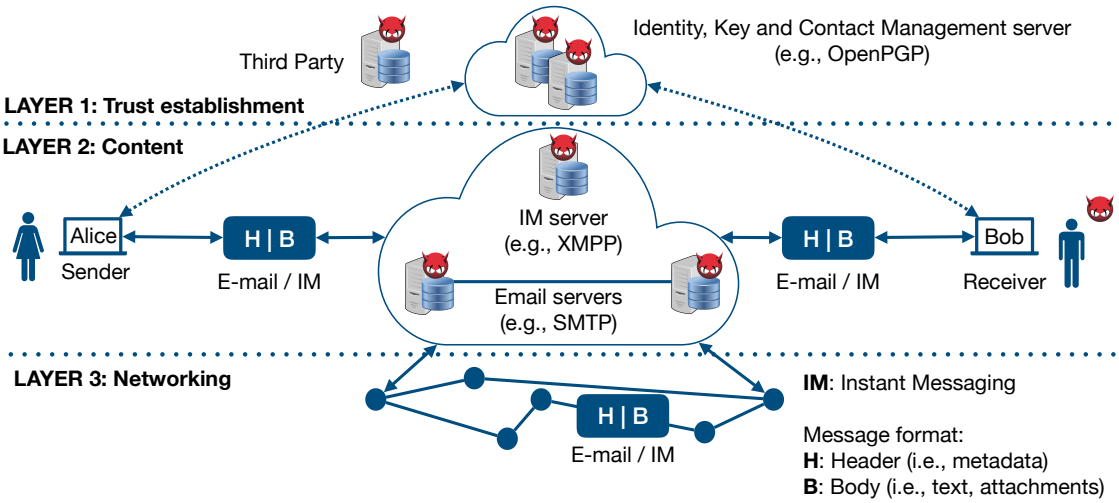

Figure 1: Communication systems: emailing, instant messaging and key management servers.

communication protocol. An *active* attacker can tamper with a protocol's messages: it can read and write messages, can execute commands for gaining access to system resources, and delay or replay messages.

Internal *vs* External. An *internal* adversary can seize control of one or of multiple entities that are part of the system. It aims to extract information from a specific entity or to prevent a message from being sent. An *external* adversary can only compromise the communication channels, eavesdrop and tamper with the messages, such as performing Man-in-The-Middle (MiTM) attacks. It can also monitor and control several parts of the network, granting the adversary the ability to correlate network traffic (Murdoch and Danezis, 2005) such as performing timing attacks (Levine et al., 2004; Zhu et al., 2010).

Attackers can combine adversarial properties to increase the effectiveness and the probability of success of an attack. For instance, an external passive attacker can monitor multiple channels of a system, whereas an active internal adversary can tamper with the messages of a targeted server (*e.g.*, the E-mailing and IM servers) (Díaz et al., 2002). In this work, we focus on TE and ME, excluding attacks on networking nodes (NetND).

Assumptions. We assume that end-points are secure and authenticated. For instance, mobile devices are malware free and users need to authenticate accessing their phones.

5.2 Security Threats and Requirements

We describe the security threats and corresponding requirements in six main classes, namely: *spoofing and entity authentication*; *tampering with data and data authentication*, *repudiation and non-repudiation*; *information disclosure and confidentiality*, *denial-of-service and availability*; *elevation of privilege and authorization* (see Table 2).

Spoofing and Entity Authentication. *Spoofing* occurs when an adversary successfully impersonates the profile of a valid user, gaining improper access to the information exchanged in the system. Spoofing is a critical step in phishing, spamming and in MiTM attacks (Hu and Wang, 2018). Here, we focus on threats related to establishing End-to-End (E2E) encryption such as *public key spoofing* attacks.

Public key spoofing threat is posed when a communicating party receives and utilize cryptographic material (*i.e.*, a public key) that does not correspond to an intended user, recipient of a message in the system. For example, spoofing is when Alice receives a public key that she assumes be Bob's key when instead the key is Eve's. That can occur in the presence of an adversarial identity management server during TE. An active identity management adversary can attempt to provide counterfeit cryptographic material, such as public keys, to the intended communicating parties.

To mitigate spoofing threats, *entity authentication* mechanisms should be in place. *Public key verification* is crucial between the communicating users in private messaging systems considering untrusted key management servers. To verify that a user is the le-

gitimate owner of a public key, peer-to-peer solutions to key verification systems exist (Finney et al., 2007; Garfinkel, 1995), where users utilizing out-of-band channels to manually to the verification. Here, users can verify the fingerprints or some more usable alternatives, like trustwords (Marques et al., 2019). Peer-to-peer solutions have limitations in key verification and management (*i.e.*, distribution, synchronization, and revocation) between the communication entities, though. Other approaches are possible to ensure authentication even when servers are untrusted. Keys can be stored encrypted, for instance. Keys can be authenticated and agreed in a distributed manner with the use of multiparty computation.

Tampering with Data and Data Authentication. *Tampering with data* occurs when an adversary alter the messages exchanged between the ME entities in the system. For instance, an adversary may (stealthily) change the content of an E-mailing or an instant message.

To contain this threat, *data authentication* of messages exchanged needs to be guaranteed: users should be able to verify that messages have not been modified in transit, and only verified messages should be accepted. Nowadays, application encryption and data authentication exist between the E-mailing and IM servers during the ME as a minimum security requirement in commercial messaging systems such as Gmail and Facebook messenger. For instance STARTTTLS offers e-mail encryption from IM (Saint-Andre, 2004) and e-mail for SMTP (Hoffman, 2002) and IMAP/POP3 (Newman, 1999). Thus, network nodes cannot read and alter the messages during message transit but the E-mailing and IM servers. Thus, end-to-end data authentication should be guaranteed such as with Message Authentication Code (MAC) and digital signatures in combination.

Repudiation, Non-repudiation, and Accountability. Adversaries can *repudiate* the status of a message to users or the ME entities of the system. For instance, an adversary may attempt to deny having sent or received an E-mailing or an instant message. Users who are involved during the communication can be considered as adversaries assuming TE and ME entities as passive adversaries. That is an reasonable assumption, as it is against their interest of companies such as Google and Facebook messenger to not perform the operations of E-mailing and IM properly.

To mitigate repudiation threats non-repudiation of actions performed and *accountability* must be guaranteed. Non-repudiation of an action can consist of a proof of *origin*, *submission*, *delivery*, and *receipt* between the intended users (Zhou and Gollmann, 1997). Non-repudiation can be achieved with the use of cryptographic primitives, such as digital signatures, and audit trails, such as timestamps.

Information Disclosure and Confidentiality. *Information disclosure* happens when an adversary succeeds in revealing the content of messages exchanged. The adversary can attempt to perform MiTM and eavesdrop the conversations of users in messaging systems. Therefore, *confidentiality of messages exchanged* within a system should be guaranteed. Communication confidentiality ensures that the messages exchanged are only readable by the involved users. Considering untrusted ME servers, the solution of E2E encryption is essential in contrast to application layer encryption aiming to mitigate MiTM attacks. Confidentiality of messages for E2E encryption systems can be achieved with encryption schemes such as symmetric, asymmetric, and malleable encryption (Borisov et al., 2004).

Another challenged to confidentiality is the storage and archiving of messages exchange. An adversary can passively collect and store messages in an attempt to retrieve or break the E2E encryption keys and compromise future or past messages. *Perfect forward* and *backward secrecy* needs to be guaranteed for private messaging. It ensures that past messages cannot be recovered retroactively. Essentially, in key agreement protocols, past session keys are not compromised even if long-term keys are compromised.

Here, we stress, there are difference between E-mailing's and IM's solutions to achieve perfect forward secrecy. For low latency protocols, *i.e.*, IM, where users are online, *perfect forward secrecy* can be achieved generating session keys such as using Diffie-Hellman key agreement (Menezes et al., 1996; Diffie and Hellman, 1976) in Off-The-Record (OTR) messaging (Borisov et al., 2004). However, the high-latency of E-mailing makes the session keys approach impractical (Borisov et al., 2004). Both Alice and Bob required to be online since the key exchange must be completed before the message is sent. *Perfect forward secrecy* for high-latency systems can be achieved with solutions such as ring signatures (Rivest et al., 2001).

If for a compromise of long-term keys subsequent cipher-texts cannot be decrypted, then the protocol guarantees also *backward secrecy* for private messaging (Borisov et al., 2004).

Denial of Service Attacks and Availability. *Denial-of-Service attack (DoS)* occurs when adversary manages to make system inaccessible to one or

more users. For instance, an adversary may target the TE and ME entities of the messaging system such as the Identity Management, the IM and the E-mailing server (Dkg, 2019; Rijhansen, 2019). The capabilities of an adversary can be from script kiddies to even state-sponsored adversaries, considered as active external TP adversaries. To reduce the threat of DoS, *availability* of the system must be guaranteed. Network security tools such as intrusion prevention systems and redundant systems such as utilizing multiparty computation are of help for this goal.

Elevation of Privilege and Authorization. *Elevation of privileges* occurs when an adversary aims to gain access to the assets of other users or the resources of the system. For instance, an adversary may attempt to become an administrator of a message group or a superuser of the system aiming at retrieving users' messages or executing operations as a super user. Therefore, *authorization* mechanisms such as access control lists that comply with the principle of least privilege for user accounts and processes should be applied.

5.3 Privacy Threats and Requirements

We describe the privacy threats and the corresponding requirements in seven main classes, namely: *linkability and unlinkability*, *identifiability and anonymity*, *non-repudiation and deniability*, *detectability and undetectability*, *information disclosure and confidentiality*, *privacy interdependence and privacy independence*, *policy noncompliance and policy compliance* (see Table 2).

Identifiability and Anonymity. *Identifiability* is defined as the extent to which a specific user can be identified from a set of users, the identifiability set. Identification is the process of linking information to allow the inference of the identity of a specific user (Cooper et al., 2013). An adversary can identify a specific user by examining an Item of Interest (IOI) such as the identity of a sender / receiver (*i.e.*, e-mail or phone number) or an action performed. For instance, a passive ME internal adversary such as the IM and E-mailing servers, may identify the sender of a message by examining the headers (*i.e.*, metadata) or the body of a message exchanged. Moreover, an active TP external adversary might try to examine the metadata of a message in transit or to correlate traffic traffic (Zhu et al., 2010) executing timing attacks (Levine et al., 2004), even in the E2E encryption messaging systems (Levine et al., 2004).

To mitigate identifiability threats, the *anonymity* of users must be guaranteed. *Anonymity* is defined from the attackers perspective as the "attacker cannot sufficiently identify the subject within a set of subjects, the anonymity set" (Pfitzmann and Hansen, 2010). In order to make anonymity possible, there always need to be a set of possible users such that for an adversary the communicating user is equally likely to be of any other user in the set (Díaz et al., 2002; Symeonidis and Hoeneisen, 2019). Thus, an adversary cannot identify who is the sender of a message. Anonymity can be achieved with the use of pseudonyms and cryptographic schemes such as anonymous re-mailers (*i.e.*, mixnets) (Shirazi et al., 2018), and secret sharing.

Linkability and Unlinkability. *Linkability* occurs when an adversary can sufficiently distinguish within a given system that two or more IOIs, such as subjects (*i.e.*, users), objects (*i.e.*, messages), or actions are related to each other (Pfitzmann and Hansen, 2010). For instance, an adversary may be able to relate pseudonyms by analysing exchanged messages and deduce that the pseudonyms belong to one user (though the user may not necessarily be identified in this process). Therefore, *unlinkability* of IOIs should be guaranteed through the use of pseudonyms as well as cryptographic schemes such as anonymous credentials (Camenisch and Lysyanskaya, 2004).

Non-repudiation and Deniability. *Non-repudiation* can be a threat to a user's privacy for private messaging systems, in contrast to security. As we prior discussed, non-repudiation should be guaranteed for users. However, non-repudiation carries a potential threat in itself when it is used against a user in certain instances. For example, whistle-blowers may find non-repudiation used against them by adversaries, particularly in countries with strict censorship policies and in cases where human lives are at stake. Adversaries in these situations may seek to use shreds of evidence collected within a communication system to prove to others that a whistle-blowing user was the originator of a specific message (Symeonidis and Hoeneisen, 2019). Therefore, *plausible deniability* is essential for these users, to ensure that an adversary can neither confirm nor contradict that a specific user sent a particular message. For example, Bob or Eve is unable to claim authorship of a message that was from Alice. Deniability can be guaranteed through the use of cryptographic protocols such as OTR for private messaging (Borisov et al., 2004).

Detectability / Observatility and Undetectability. *Detectability* occurs when an adversary is able to sufficiently distinguish an IOI, such as messages exchanged within the system, from random noise (Pfitzmann and Hansen, 2010). Observability occurs when that detectability occurs along with a loss of anonymity for the entities within that same system. An adversary can exploit the states of detectability and anonymity in order to detect, infer and possibly identify users within a system. The threat posed by an adversary can be a passive ME internal adversary eavesdropping communication of messages exchanged. It can also be an active TP external by seizing control of several entities and communication links. That grants the adversary the ability to correlate and control traffic (Zhu et al., 2010) in order to execute timing attacks, even in the end-to-end communication systems (Levine et al., 2004).

Therefore, *undetectability* of IOIs should be guaranteed, which together with preserving the anonymity of users it also ensures unobservability. Undetectability for an IOI is defined as that "the attacker cannot sufficiently distinguish whether it exists or not" while unobservability also incorporates "anonymity of the subject(s) involved in the IOI" (Pfitzmann and Hansen, 2010). Undetectability can be achieved through the use of cryptographic schemes such as mix-nets (Shirazi et al., 2018; Symeonidis and Hoeneisen, 2019) and obfuscation mechanisms such as the insertion of dummy traffic within a system.

Information Disclosure and Confidentiality. *Information disclosure* - or loss of confidentiality - about users, message content, metadata, or other information is not only a security but also a privacy threat that a communicating system can face. For example, a successful MiTM attack can yield message content and metadata that can be used to determine not only the content of a message by also with whom a specific user communicates with, and how frequently. To guarantee the *confidentiality* of messages and prevent information disclosure, security measures need to be guaranteed with the use of cryptographic primitives and protocols such as symmetric, asymmetric or homomorphic encryption and secret sharing.

Privacy Interdependence and Privacy Independence. *Privacy interdependence* is defined as the scenario when the privacy of users is affected by decisions taken and actions initiated by anyone but the user thyself (Symeonidis, 2018). Privacy interdependence is related to relational and spatial privacy as there were defined by Clark (Roger Clarke, 2016) and Biczók (Biczók and Chia, 2013). Relational privacy considers how a user relates to and communicate with other users (Biczók and Chia, 2013). For instance, a passive TE (i.e., identity, and key management server) and ME servers can learn the contacts that a user is related to such as names, phone numbers, E-mailing addresses and public keys from the contacts list of a user. With *spatial privacy* a user is affected by the actions of another user on the invasion on the virtual online space of a user (Biczók and Chia, 2013). For instance, a user might forward a message to other users without the notification and consent of the involved users affecting their privacy. Therefore, *independent privacy* is essential and must be guaranteed for each user in private messaging systems. It needs to ensure the transparency and direct consent to the information sharing by the users that are involved in communicating. Moreover, to safeguard *individual privacy* protocols considering cryptographic primitives such as zero-knowledge proofs and multiparty computation needs to be designed and implemented.

Policy Non-compliance and Policy Compliance *Policy non-compliance* can be a threat to the privacy of users in a private messaging system. An adversary, can attempt to collect and process information about users in E-mailing and IM without the users' awareness, notification and explicit consent such as for advertisement purposes. That can result in unauthorized processing of users information under the GDPR (European Parliament and the European Council, 2016) resulting in profiling, and censorship. Therefore, data protection *policy compliance* must be guaranteed. It can be achieved with auditing tools such as with Data Protection Impact Assessment (European Commission,) considering the (European Parliament and the European Council, 2016).

6 LIMITATIONS AND DISCUSSION

We compiled a preliminary, yet rich, list threats and requirements for private messaging systems. However, to fulfill the requirements, one needs to refer to specific cryptographic protocols. This task is left as future work.

At the level where our work stands (*i.e.*, threats and requirements) certain system differences between E-mailing and IM are neglected. We expect these differences emerging and impacting design choices. For example, OTRv4 protocol (Bini and Celi, 2018) considers these difference to provide E2E while preserving the same security and privacy requirements for both systems.

7 CONCLUSION

This paper's goal has been that of systematically discuss threats against private messaging in the two domains of E-mailing and IM when servers are untrusted. The two application types have strong similarities, which allowed us to apply for both the same threat modelling methodologies; but they are not identical and already from our analysis relevant difference emerged due to the different nature, on-line and off-line, of the two communication paradigms. Fulfil the same requirements, such as perfect forward secrecy, calls for different solutions.

We provided a detailed and comparative analysis of the systems, describing the entities involved, the operations, and the system features. We have identified, discussed, and catalogued the classes of threats that can exploit the assets of users and the system. We have, for each threat, discussed the opposing security and privacy requirements.

Our investigation remains within the scope of specific features of privacy messaging: messages exchange, search and archiving, and contact synchronization. We intend for the future to extend the analysis to other features, such as, key management and key synchronization among different devices.

ACKNOWLEDGMENTS

We would like to thank Nana Karlstetter and Bernie Hoeneisen from p≡p foundation for the inputs and all the anonymous reviewers for their constructive feedback. The authors are supported by the pEp Security SA / SnT partnership project "Protocols for Privacy Security Analysis".

REFERENCES

Biczók, G. and Chia, P. H. (2013). Interdependent privacy: Let me share your data. In *Financial Cryptography and Data Security - 17th International Conference, FC 2013, Okinawa, Japan, April 1-5, 2013, Revised Selected Papers*, pages 338–353.

Bini, O. and Celi, S. (2018). No evidence of communication and implementing a protocol: Off-the-record protocol version 4. *Hotpets*.

Borisov, N., Goldberg, I., and Brewer, E. A. (2004). Off-the-record communication, or, why not to use PGP. In *Proceedings of the 2004 ACM Workshop on Privacy in the Electronic Society, WPES 2004, Washington, DC, USA, October 28, 2004*, pages 77–84.

Camenisch, J. and Lysyanskaya, A. (2004). Signature schemes and anonymous credentials from bilinear maps. In *Advances in Cryptology - CRYPTO 2004, 24th Annual International CryptologyConference, Santa Barbara, California, USA, August 15-19, 2004, Proceedings*, pages 56–72.

Clark, J., van Oorschot, P. C., Ruoti, S., Seamons, K. E., and Zappala, D. (2018). Securing email. *CoRR*, abs/1804.07706.

Cooper, A., Tschofenig, H., Aboba, B., Peterson, J., Morris, J., Hansen, M., and Smith, R. (2013). Privacy Considerations for Internet Protocols. RFC 6973.

Council of Europe: European Court of Human Rights (2016). Guide on Article 8 of the European Convention on Human Rights - Right to respect for private and family life. https://www.echr.coe.int/Documents/Guide_Art_8_ENG.pdf. Accessed December, 2019.

Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J., Métayer, D. L., Tirtea, R., and Schiffner, S. (2015). Privacy and data protection by design - from policy to engineering. *CoRR*, abs/1501.03726.

Deng, M., Wuyts, K., Scandariato, R., Preneel, B., and Joosen, W. (2011). A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. *Requirements Engineering*, 16(1):3–32.

Díaz, C., Seys, S., Claessens, J., and Preneel, B. (2002). Towards measuring anonymity. In *Privacy Enhancing Technologies, Second International Workshop, PET 2002, San Francisco, CA, USA, April 14-15, 2002, Revised Papers*, pages 54–68.

Diffie, W. and Hellman, M. E. (1976). New directions in cryptography. *IEEE Trans. Information Theory*, 22(6):644–654.

Dkg (2019). OpenPGP Certificate Flooding. https://dkg.fifthhorseman.net/blog/openpgp-certificate-flooding.html. Accessed December, 2019.

Ermoshina, K., Musiani, F., and Halpin, H. (2016). End-to-end encrypted messaging protocols: An overview. In *Internet Science - Third International Conference, INSCI 2016, Florence, Italy, September 12-14, 2016, Proceedings*, pages 244–254.

European Commission. Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679 (17/EN WP 248). http://ec.europa.eu/newsroom/document.cfm?doc_id=44137. Accessed May, 2018.

European Parliament and the European Council (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L119, 4.5.2016. 59:1–87.

Finney, H., Donnerhacke, L., Callas, J., Thayer, R. L., and Shaw, D. (2007). OpenPGP Message Format. RFC 4880.

Garfinkel, S. (1995). *PGP - pretty good privacy: encryption for everyone (2. ed.)*. O'Reilly.

Hoffman, P. (2002). SMTP Service Extension for Secure SMTP over Transport Layer Security. RFC 3207.

Howard, M. and Lipner, S. (2009). *The security development lifecycle*. O'Reilly Media, Incorporated.

Hu, H. and Wang, G. (2018). End-to-end measurements of email spoofing attacks. In *27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018.*, pages 1095–1112.

Klensin, J. (2008). Simple Mail Transfer Protocol. RFC 5321.

Levine, B., Reiter, M., Wang, C., and Wright, M. (2004). Timing attacks in low-latency mix systems (extended abstract). In *Financial Cryptography, 8th International Conference, FC 2004, Key West, FL, USA, February 9-12, 2004. Revised Papers*, pages 251–265.

Marotta, D. J. and Russell, M. (2013). Support for The Right to Privacy of Correspondence. http://www.marottaonmoney.com/support-for-the-right-to-privacy-of-correspondence/. Accessed December, 2019.

Marques, H., Luck, C., and Hoeneisen, B. (2019). pretty Easy privacy (pEp): Privacy by Default. Internet-Draft draft-birk-pep-04, Internet Engineering Task Force. Work in Progress.

Menezes, A., Oorschot, P., and Vanstone, S. (1996). *Handbook of Applied Cryptography*. CRC Press.

Microsoft (2010). Improving Web Application Security: Threats and Countermeasures. Accessed May, 2016.

Murdoch, S. J. and Danezis, G. (2005). Low-cost traffic analysis of tor. In *2005 IEEE Symposium on Security and Privacy (S&P 2005), 8-11 May 2005, Oakland, CA, USA*, pages 183–195.

Newman, C. (1999). Using TLS with IMAP, POP3 and ACAP. RFC 2595.

Pfitzmann, A. and Hansen, M. (2010). A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management.

Resnick, P. (2008). Internet Message Format. RFC 5322.

Rijhansen (2019). SKS Keyserver Network Under Attack. https://gist.github.com/rjhansen. Accessed December, 2019.

Rivest, R. L., Shamir, A., and Tauman, Y. (2001). How to leak a secret. In *Advances in Cryptology - ASIACRYPT 2001, 7th International Conference on the Theory and Application of Cryptology and Information Security, Gold Coast, Australia, December 9-13, 2001, Proceedings*, pages 552–565.

Roger Clarke (2016). Introduction to Dataveillance and Information Privacy, and Definitions of Terms (1997) (revised in 1999, 2005, 2006). http://www.rogerclarke.com/DV/Intro.html. Accessed October, 2019.

Saint-Andre, P. (2004). Extensible Messaging and Presence Protocol (XMPP): Core. RFC 3920.

Saint-Andre, P. (2011). Extensible Messaging and Presence Protocol (XMPP): Instant Messaging and Presence. RFC 6121.

Shirazi, F., Simeonovski, M., Asghar, M. R., Backes, M., and Díaz, C. (2018). A survey on routing in anonymous communication protocols. *ACM Comput. Surv.*, 51(3):51:1–51:39.

Signal (2013). Forward secrecy for asynchronous messages. https://signal.org/blog/asynchronous-security. Accessed July, 2019.

Stanger, A. (2019). *Whistleblowers: Honesty in America from Washington to Trump*. Yale Univ. Press.

Symeonidis, I. (2018). *Analysis and Design of Privacy-Enhancing Information Sharing Systems*. PhD thesis, ESAT, imec-COSIC, KU Leuven , Kasteelpark Arenberg 10, 3001 Leuven, Belgium. http://hdl.handle.net/10993/37607.

Symeonidis, I., Aly, A., Mustafa, M. A., Mennink, B., Dhooghe, S., and Preneel, B. (2017). Sepcar: A secure and privacy-enhancing protocol for car access provision. In *Computer Security - ESORICS 2017 - 22nd European Symposium on Research in Computer Security, Oslo, Norway, September 11-15, 2017, Proceedings, Part II*, pages 475–493.

Symeonidis, I. and Hoeneisen, B. (2019). Privacy and Security Threat Analysis and Requirements for Private Messaging. Internet-Draft draft-symeonidis-medup-requirements-00, Internet Engineering Task Force. Work in Progress.

Unger, N., Dechand, S., Bonneau, J., Fahl, S., Perl, H., Goldberg, I., and Smith, M. (2015). Sok: Secure messaging. In *2015 IEEE Symposium on Security and Privacy, SP 2015, San Jose, CA, USA, May 17-21, 2015*, pages 232–249.

Zhou, J. and Gollmann, D. (1997). Evidence and nonrepudiation. *J. Netw. Comput. Appl.*, 20(3):267–281.

Zhu, Y., Fu, X., Graham, B., Bettati, R., and Zhao, W. (2010). Correlation-based traffic analysis attacks on anonymity networks. *IEEE Trans. Parallel Distrib. Syst.*, 21(7):954–967.

Table 2: Entities that can be utilized as adversaries for private messaging systems.

Adversarial model: **Passive □ / Active ⊠ — Internal ⊂ (TE, ME, User)** **Passive □ / Active ⊠ — External ≍ (TP)**				**User Assets**				**System Assets**			
				ID **(e-mail, phone #)** **(Pk, certificates)**	**Contacts** **list**	**Message content** **(Text, attachments)**	**Metadata** **(Sender, receiver, packet size)**	**TE Servers** **ME Servers**			
Classes of Threats									**Classes of Requirements**		
S	**Spoofing**		S	⊠ — ⊂ (TE) ⊠ — ≍ (TP)	-	-	-	-	EA	**Entity Authentication**	
S	**Tampering with Data**		T	-	-	⊠ — ⊂ (ME) ⊠ — ≍ (TP)	-	-	DA	**Data Authentication**	
S	**Repudiation**	**Origin / Delivery**	R	-	-	⊠ — ⊂ (User)	-	-	NR	**Origin / Delivery**	**Non-repudiation**
S	**Repudiation**	**Submission / Receipt**	R	-	-	⊠ — ⊂ (User)	-	-	NR	**Submission / Receipt**	**Non-repudiation**
S	**Information Disclosure**	**Current**	I	-	-	⊠ — ⊂ (ME) ⊠ — ≍ (TP)	-	-	C	**Current**	**Confidentiality**
S	**Information Disclosure**	**Previous / Future Messages**	I	-	-	⊠ — ⊂ (ME) ⊠ — ≍ (TP)	-	-	C	**Forward / Backward Secrecy**	**Confidentiality**
S	**Denial of Service**		D	-	-	-	-	⊠ — ≍ (TP)	Av	**Availability**	
S	**Elevation of Privilege**		E	-	-	-	-	⊠ — ⊂ (User) ⊠ — ≍ (TP)	A	**Authorization**	
P	**Linkability**		L	-	-	□ — ⊂ (ME) ⊠ — ≍ (TP)	□ — ⊂ (ME) ⊠ — ≍ (TP)	-	An	**Unlinkability**	
P	**Identifiability**		I	-	-	□ — ⊂ (ME) ⊠ — ≍ (TP)	□ — ⊂ (ME) ⊠ — ≍ (TP)	-	Un	**Anonymity**	
P	**Non-repudiation**	**Origin / Delivery**	N	-	-	□ — ⊂ (ME, User(receiver)) ⊠ — ≍ (TP)	□ — ⊂ (ME) ⊠ — ≍ (TP)	-	Den	**Origin / Delivery**	**Deniability**
P	**Non-repudiation**	**Submission / Receipt**	N	-	-	□ — ⊂ (ME, User(receiver)) ⊠ — ≍ (TP)	□ — ⊂ (ME) ⊠ — ≍ (TP)	-	Den	**Submission / Receipt**	**Deniability**
P	**Detectability**		D	-	-	□ — ⊂ (ME) ⊠ — ≍ (TP)	□ — ⊂ (ME) ⊠ — ≍ (TP)	-	UnDect	**Undetectability**	
P	**Information Disclosure**	Current	D	-	-	□ — ⊂ (ME) ⊠ — ≍ (TP)	□ — ⊂ (ME) ⊠ — ≍ (TP)	-	C	**Current**	**Confidentiality**
P	**Information Disclosure**	**Previous / Future**	D	-	-	□ — ⊂ (ME) ⊠ — ≍ (TP)	□ — ⊂ (ME) ⊠ — ≍ (TP)	-	C	**Forward / Backward**	**Confidentiality**
P	**Privacy Interdependence**		I	-	□ — ⊂ (TE, ME, User)	□ — ⊂ (ME, User(receiver))	-	-	PI	**Privacy Independence**	
P	**Policy non-compliance**		N	□ — ⊂ (TE, ME)	□ — ⊂ (TE, ME)	□ — ⊂ (TE, ME)	□ — ⊂ (TE, ME)	-	PN	**Policy compliance**	