27 Years and 81 Million Opportunities Later: Investigating the Use of Email Encryption for an Entire University

Christian Stransky*, Oliver Wiese†, Volker Roth†, Yasemin Acar‡, and Sascha Fahl*§
*Leibniz University Hannover, stransky@sec.uni-hannover.de
†Freie Universität Berlin, {oliver.wiese, volker.roth}@fu-berlin.de
‡Max Planck Institute for Security and Privacy, yasemin.acar@mpi-sp.org
§CISPA Helmholtz-Center for Information Security, sascha.fahl@cispa.de

***Abstract*—Email is one of the main communication tools and has seen significant adoption in the past decades. However, emails are sent in plain text by default and allow attackers easy access. Users can protect their emails by end-to-end encrypting them using tools such as S/MIME or PGP.**

Although PGP had already been introduced in 1991, it is a commonly held belief that email encryption is a niche tool that has not seen widespread adoption to date. Previous user studies identified ample usability issues with email encryption such as key management and user interface challenges, which likely contribute to the limited success of email encryption.

However, so far ground truth based on longitudinal field data is missing in the literature. Towards filling this gap, we measure the use of email encryption based on 27 years of data for 37,089 users at a large university. While attending to ethical and data privacy concerns, we were able to analyze the use of S/MIME and PGP in 81,612,595 emails.

We found that only 5.46% of all users ever used S/MIME or PGP. This led to 0.06% encrypted and 2.8% signed emails. Users were more likely to use S/MIME than PGP by a factor of six. We saw that using multiple email clients had a negative impact on signing as well as encrypting emails and that only 3.36% of all emails between S/MIME users who had previously exchanged certificates were encrypted on average.

Our results imply that the adoption of email encryption is indeed very low and that key management challenges negatively impact even users who have set up S/MIME or PGP previously.

I. Introduction

Email is one of the major online communication tools. As of February 2021, there are more than 4 billion email users worldwide sending and receiving over 300 billion emails per day [41]. While email is used for all kinds of information including the most sensitive kinds such as trade secrets, account credentials, and health data, regular email is not encrypted and allows network attackers and service providers unauthorized access. This is not for a lack of tools. Both S/MIME [14] and PGP [46] were introduced almost 30 years ago with the goal to provide end-to-end encryption for email. However, in contrast to modern messaging tools such as Signal [37] or WhatsApp [43] that implement end-to-end encryption by default, S/MIME and PGP require a complex manual setup by users. Consequently, previous work has shown that using email encryption correctly and securely is challenging for many users [20], [32], [34]–[36], [45]. They struggle with setting up and configuring encryption keys, distributing them, managing keys on multiple devices, and revoking them. These findings, already anticipated by Davis [12], are corroborated by public reports of failed PGP use. For example, it took Edward Snowden and the journalist Glenn Greenwald a few months and serious effort to set up PGP for email in order to communicate securely [28]. Hence, it is commonly believed in the security community that end-to-end encrypted email is not widely used, mostly because of lacking usability and awareness issues identified in a multitude of user studies in the past 22 years (cf. [11], [20]–[22], [30], [34], [36]). To the best of our knowledge, our work is the first scientific collection and evaluation of ground truth on the adoption of end-to-end email encryption. Our work is mainly motivated as follows:

Ground Truth. We aim to confirm the security community's anecdotal knowledge about the low adoption of end-to-end email encryption and provide ground truth based on field data. Our longitudinal field data can help motivate future work to improve the adoption of end-to-end encryption for email.

Method Extension. We extend the toolbox of the past 22 years of email encryption research that was initiated with the seminal paper "Why Johnny Can't Encrypt" [45] at USENIX Security'99 that is mostly based on laboratory experiments and self-reporting studies: In this work, we investigate a large dataset including millions of data points of thousands of users and years of their email data.

Validate Results from Previous Work and Investigate Underexplored Challenges. We confirm findings from previous work (e.g. [1], [26], [27], [30]) obtained by other methods including smaller-scale interviews, surveys, and controlled experiments. Additionally, we also investigate further challenges that require large scale field data.

Motivated by the above, we make the following contributions in the course of this work:

Data Collection Pipeline. In collaboration with our data protection officer, university staff council, and the technical staff of the university IT department, we developed and tested

a reproducible and privacy friendly data collection pipeline that allows to analyze large amounts of email data with a focus on S/MIME and PGP usage (cf. Section IV-A). The data collection pipeline is part of our replication package. We aim to encourage other institutions to investigate their adoption of S/MIME and PGP to contribute to an even better understanding of the email encryption ecosystem.

Adoption of Email Encryption at a Large University. We provide a detailed evaluation of the adoption of email encryption at our university in the past 27 years. In our evaluation, we focus on the use of S/MIME and PGP for 37,089 total email accounts. Our investigation of 81,612,595 emails found that 2.8% of them were digitally signed and 0.06% were encrypted. We found that only 5.46% of our users ever used S/MIME or PGP and that S/MIME was more widely used than PGP. However, PGP was the more popular email encryption tool amongst researchers.

Use of S/MIME and PGP. We provide a detailed overview of S/MIME certificates and PGP keys in our dataset and find that RSA is the most widely used key algorithm, employing 2048 bits keys most often for S/MIME. PGP keys used 4096 bits most often, although newer PGP keys used less secure 2048 bits. We find that more than one third of all PGP keys did not have an expiration date set making revocation unnecessarily complicated and *Deutsche Telekom* to be the root CA for 64.95% of all S/MIME certificates.

User Interaction Challenges including Key Management. We report on an investigation of user interaction challenges that previous work identified in user studies. Most interestingly, we focus on key management issues during key distribution, multi device use, and key rollover. We find that even after exchanging public keys, only 3.36% of all emails between S/MIME users were encrypted on average. The use of multiple email clients had a negative impact on the amount of signed and encrypted emails. While most S/MIME and PGP users renewed their keys in time, 11.5% of S/MIME key rollovers occurred after the keys' expiration.

Overall, our results confirm the pessimistic assessment of the security community: Although our university provides all researchers, staff, and students with free access to S/MIME certificates, only very few make use of them and only a negligible amount of emails was encrypted or signed. Our findings also support results from previous user studies and illustrate additional challenges. Management of email encryption keys is hard and distributing keys, using multiple email clients, or having to renew keys complicates matters.

The rest of the paper[1] is organized as follows: In Section II we provide information on S/MIME, PGP, and our university's S/MIME certificate authority. We provide an overview of related work and contextualize our contributions in Section III. In Section IV, we describe our methodology by providing details for our data collection pipeline, discussing ethical and data privacy implications of our work, illustrating limitations, and summarizing the replication package. In Section V, we provide detailed results of our evaluation, discuss their implications in Section VI, and conclude the paper in Section VII.

II. Background

In this section, we provide background information on OpenPGP, S/MIME, and the email ecosystem of our university including its S/MIME certificate authority.

A. OpenPGP

OpenPGP[2] is an encryption standard (cf. [10], [15]) which is used for email encryption and digital signatures. PGP is an open source project and was first standardized in 1996. In the first standardization, PGP messages were added to the text body (named: PGP Inline) of an email. Later versions introduced a separate MIME type for PGP messages (named: PGP MIME). Over time, new algorithms have been added, including the Camellia and ECDSA cryptography algorithms. PGP supports the use of key servers for public key exchange. Users can search these servers for keys for given email addresses. However, keys may also be exchanged by attaching public keys to emails. Additionally, several email clients, like K9 on Android or Thunderbird using the Enigma plugin, support hidden key exchange by adding public keys to email headers. This feature has been standardized and further developed by the open source project Autocrypt[3] since 2016.

In contrast to centralised trust infrastructures known from the web PKI or S/MIME, PGP relies on the Web of Trust to verify identities. In the web of trust approach, users sign each other's key when meeting other PGP users in person. Therefore, users can trust a new key if another trusted key previously signed the new key, relying on a decentralized trust chain.

B. S/MIME

Secure/Multipurpose Internet Mail Extensions (S/MIME) is a standard to encrypt and sign emails. It was first introduced in 1998 (RFC2311 [15]) and has constantly been improved since then. S/MIME utilizes a Public Key Infrastructure to verify certificates and as such has mostly been used in corporate environments, where a certificate authority (CA) is deployed or third party CAs are utilized to issue certificates to employees. It has been widely supported out-of-the box without the need for third party plugins in commercially used email clients like Outlook 98 and higher or Thunderbird.

C. Email Ecosystem at our University

Email at our university is a centralized service. The university's computing center provides email accounts for all administrative staff members, for all students as well as faculties, departments, and research groups. Overall, our university offers 90 different study subjects reaching from engineering to humanities, and has about 30,000 students and 5,000

[1] Find our companion website at: https://publications.teamusec.de/2022-oakland-email/

[2] Abbreviated as PGP in the paper

[3] cf. https://autocrypt.org

staff members. Faculties, departments and research groups are organized in decentralized units, i.e. each faculty, and most departments and research groups have their own subdomain (e.g., sec.uni-hannover.de for the information security research group) for their email. Users can access their email accounts either through a web interface or dedicated email clients using the university's POP3, or IMAP and SMTP servers.

D. University Certificate Authority / Registration Authority

Our university is part of the public key infrastructure of the communications network for science and research in Germany (DFN).[4] The university's computing center provides a registration authority for the DFN CA to issue certificates for email end-to-end encryption and signing, server authentication for TLS, and document signing for its scientific and administrative staff and all students (DFN-PKI).[5] Certificate signing and revocation is processed through the DFN CA.

Certificate Policies. All university employees and students are eligible to obtain S/MIME certificates. However, certificate use is neither officially endorsed, nor are issuances automatically triggered. Individual work groups may informally encourage certificate use. While the CA also provides server certificates (e.g., for TLS), our work focuses on user certificates for email encryption/signing.

User Certificates. To apply for a certificate, eligible staff and students can apply online, receive a certificate signing request, make an appointment with the registration authority, show up in person, present a proof of identity, and then receive a valid certificate. This process is comparably complicated. In contrast, creation of a student ID that can be used to access free transport and student discounts, and is used as proof of identity in exams, does not require in-person interaction. The process is also not embedded in any other existing onboarding process at the university. New certificates for the same user can be issued without another identification if the last identification is not older than 39 months.

Certificate Signing Request Process. Users can generate a certificate signing request (CSR) for email certificates using a web application entering their personal details and a revocation pin. The web application generates a CSR, saves it in the browser certificate store, and asks the user to enter a passphrase which protects the CSR's private key.

Certificate Revocation Process. The CA provides a web interface for certificate revocation. To revoke certificates, users have to enter their certificate's serial number and a revocation PIN.

Certificate Expiration. Once 30 days and a second time 15 days before the user certificate expires, the DFN sends users an expiration warning via email and encourages users to renew the certificate.

History and Milestones. Our university began to issue S/MIME certificates in 2004. The first root certificate used to sign the original CA (G1) through the DFN expired in 2019. Starting from 2017, certificates were issued using a new root CA (G2). Additionally, 479 certificates issued using the G1 CA had to be replaced with new certificates issued by the G2 root CA. Figure 10 in Appendix C illustrates the issued certificates per year. In the first years, the amount of server and user certificates is roughly equal. However, the number of S/MIME certificates is only slowly increasing compared to the number of server certificates. We can see a drop in requested certificates in 2020, probably due to the COVID-19 pandemic.

III. Related Work

We discuss related work in two key areas and put our work into context: User studies for end-to-end encryption with end users and email field studies.

A. End-to-End Encryption Studies for End Users

The usability of end-to-end encryption has been a research focus for decades. With their seminal work, Whitten and Tygar [45] set this line of research off in 1999 when they evaluated the usability of PGP and identified challenges for end-users in a lab study. In their qualitative user study, participants had trouble with key management, specifically key exchange with other participants. One participant forgot the password for their key pair and thus had to generate a new one. Another participant was unable to encrypt an email and others were unable to decrypt emails. Following Whitten and Tygar, Garfinkel and Miller [20] and Roth [29] hypothesized that some of the challenges could be overcome by simplifying the key verification process. Their solutions were based on omitting the verification of keys by third parties and instead using a trust on first use (TOFU) approach. Garfinkel and Miller tested their approach on lay users and found that color-coding messages depending on their signature status makes users significantly less susceptible to social engineering attacks overall. Garfinkel et al. surveyed 470 merchants who received digitally-signed VAT invoices from Amazon and found that merchants should send signed emails by default as the passive exposure seems to increase acceptance and trust [19], [21]. Ruoti et al. studied different self-designed and publicly available encryption tools [31]–[33], [35] to improve the usability of email encryption in their laboratory. Their research interests were key management, key distribution, and automatically vs. manually enabled encryption. In several laboratory studies by Ruoti et al. [31], [32], [35], participants rated at least one tool as usable and indicated interest in secure email, but did not know when or how they wanted to use it. In another recent journal article, Ruoti et al. came to the conclusion that secure and usable systems so far had only been tested in short-term studies and future research should investigate long term usability and adoptability of secure email systems [30]. Atwater et al. and Lerner et el. proposed clients for PGP similar to Keybase[6] to study how to simplify key distribution [5], [25].

[4]https://www.dfn.de/en/

[5]https://www.pki.dfn.de/ueberblick-dfn-pki/ (german)

[6]https://keybase.io/

They proposed to upload users' public keys to a website and confirm ownership such that they can retrieve emails from the corresponding email address. In a lab study, Atwater et al. found that such a key distribution mechanism enabled participants to send more encrypted emails and had improved usability for (webmail) users. Lerner et al. compared their tool called Confidante with Mailvelope[7] and showed that their tool reduces the error potential [25]. In a lab study, all participants in the Mailvelope group struggled to import public keys or to share their own public key. In the Confidante group, three out of nine participants struggled to import public keys. However, all of them managed to share their public keys successfully. Bai et al. proposed encryption prototypes to study user flows of different key management approaches [8], [9]. In an interview study, participants preferred to register their public keys on a webpage and automatically retrieve public keys of communication partners from the webpage over manual key management. Fahl et al. examined different usability aspects for Facebook message encryption mechanisms and found that automatic key management and key recovery capabilities are important for adoption [16]. McGregor et al. reported that cooperating journalists used PGP to encrypt their emails when investigating the Panama Papers [27]. However, journalists also identified obstacles when using encryption with multiple devices. Consequently, they used secure messengers such as Signal instead of PGP on their smartphones. Gaw et al. interviewed nine employees from the same company and found that users flagged encrypted mails as urgent and found those to be annoying when used for all messages [22]. They argued that understanding of social factors is important for adoption. In a combined lab and field study, Mauriés et al. participants struggled with Enigmail for Thunderbird and the Mailvelope browser plugin [26]. Enigmail users needed help to setup the tool on their computer. The setup process in Mailevelope was unclear. One participant struggled to import a public key and send an encrypted email.

In addition to email, mobile messaging apps including Signal, Threema and WhatsApp, made end-to-end encryption available for the masses. WhatsApp, for example, introduced end-to-end encrypted messages for all its users by default in 2016 [43]. In an interview study, Abu-Salma et al. identified blockers and barriers for the adoption of end-to-end encryption including incompatible tooling and misconceptions of end-to-end encryption features [2]. They argue that usability is not the primary obstacle and that fragmented userbases or a lack of multi-device support significantly contribute to the non-adoption of end-to-end encryption. In a different study, Abu-Salma et al. further explored users' mental models and found misconceptions about security properties of messengers [1]. They argue that adoption is no longer the main challenge for end-to-end encryption tools but that people instead switch to non secure communication tools and need assistance in choosing the right one for sensitive information. Stransky et al. confirmed these findings in an online study with WhatsApp users. They found that security perceptions of end-to-end encryption in mobile messaging apps heavily depended on the reputation and expectations of an app provider, while visualizing encryption has only limited impact on perceived security [38]. Similarily, Akgul et al. found that participants noticed educational messages and that they improved understanding of security concepts when they are used in isolation. However, when those messages were implemented in a realistic environment, they could not find significant improvements in the mental models of end-to-end encryption of users [3].

Overall, previous work primarily focused on identifying blockers and barriers to adopting end-to-end encryption for email or mobile messaging apps and studied alternatives to or extensions of existing approaches using laboratory and interview studies. In contrast, we aim to extend the toolbox of end-to-end encryption research and provide ground truth based on longitudinal field data by evaluating a large dataset including millions of data points of thousands of users and years of their email data.

B. Email Field Studies

In addition to the user studies discussed above, researchers performed multiple field studies on the use of email. In 1996, Whittaker and Sidner analyzed the mailboxes of 18 NotesMail users containing 2,482 emails. They postulated email overload and studied how users handle a mass of emails [44]. In 2006, Fisher et al. revisited this analysis with a sample of 600 mailboxes containing 28,660 emails [17]. They analyzed users' email sorting strategies, especially with respect to dealing with the increased volume of emails, and postulated that large folders would make email retrieval hard. Alrashed et al. studied a sample of anonymized email logs from Outlook Web Access over a four month period, containing about 800 million actions [4]. They aimed to understand how users handle incoming emails. They find that most emails have a short lifetime and that deleting email is the most common action on messages users interacted with once. Avigdor-Elgrabli et al. analyzed a sample of donated mailboxes of a major email service provider, containing about 5 million emails [7]. They used machine learning techniques to identify relationships between emails. Roth et al. performed a study with anonymized mailboxes from 17 voluntary users that contained approximately 139,000 mails to investigate which security mechanisms would be most appropriate for their communication patterns [29]. They argue that for individual non-commercial users, out-of-band verification of keys would be more feasible than relying on public key certificates issued by third parties.

In addition to the above field studies that focused on email usability more broadly, researchers investigated the adoption of security protocols for email. Foster et al. scanned the Simple Mail Transfer Protocol (SMTP) configurations of about 300,000 major email providers and email generators in March 2014 and February 2015, and investigated the behavior of known email providers [18]. They found that TLS is widely used and discovered a dramatically low adoption of effective TLS certificate validation. Durumeric et al. examined log data

[7]Mailvelope is a browser plugin for PGP: https://www.mailvelope.com/

for SMTP handshakes of Google's Gmail service from January 2014 to April 2015 and compared it with a snapshot of SMTP configs of Alexa top million domains as of April 2015 [13]. The authors examined the distribution and use of TLS and other server-side security mechanisms. They found that the top mail providers proactively encrypted and authenticated messages. However, these practices had yet to reach widespread adoption in a long tail of over 700,000 SMTP servers with less secure configurations. Ulrich et al. evaluated the PGP key database (SKS-Keyserver) in December 2009, examining 2.7 million keys of which 400,000 were expired and 100,000 were revoked [42].

Overall, the above field studies investigate the adoption and usability of email in a broader context, measure email servers' security configurations, and conduct small-scale security analyses of email encryption. We extend previous field studies by focusing on adopting email encryption using a longitudinal large field dataset.

IV. Analyzing 27 Years of Email Data

Below, we provide detailed information on the data collection and analysis process in our work.

We performed our longitudinal analysis of a large email dataset in coordination with the technical staff of the university's IT department, the data protection officer, and the university's staff council (see Section IV-C for more details). We implemented a data collection pipeline to collect pseudonymized[8] metadata for all email accounts, including the use of S/MIME and PGP (cf. Figure 1) at our institution in the last 27 years. At no point did we collect email subject or body information to avoid the disclosure of personally identifiable information (PII) to the researchers. We also ensured that metadata included neither email account names nor the departments' names or subdomains. We aimed to keep the number of processing errors low and consistently tested the pipeline with our own mailboxes until no further processing errors occurred.

The IT department's technical staff reviewed the pipeline for functionality and data protection aspects, and then executed it on the university's standby backup email server. The backup server is a hot-standby of the primary mail server, and automatically takes over in case of a failure. Data between both servers is constantly synchronized and as a result, the backup is identical to the live data. The backup server retains all email data, dating back to early 1994.

Figure 1 provides an overview of the nine-step processing pipeline: We initially started with a local testing environment on a small sample emailbox created specifically for our study **(1)**; The technical staff reviewed the initial pipeline and iteratively tested it against the full set of emailboxes of the researchers and their own emailboxes **(2)**; We exported the emailboxes to JSON-formatted files **(3)**; We parsed and pseudonymized all emails **(4)**; We performed assertion checks on every email to ensure that neither the emailaddress nor the domain was present in any result fields to account for emailclients writing private data to unexpected places **(5)**; In the case of a succeeding assertion check, we stored the resulting email metadata for further analysis on a secure server in the university's computing center **(6a)**; In the event of a failure, we dropped all email metadata to avoid the leakage of private information **(6b)**; The IT department's technical staff transferred pseudonymized results to the authors' secure cloud storage **(7)**; We analyzed the pseudonymized results **(8)**.

[8]The pseudonymization process is described in Sections IV-A and IV-C.

A. Data Collection Pipeline

The university uses Dovecot[9] for their email servers. Dovecot offers an export feature to extract all emails as a JSON-formatted file. We implemented our processing pipeline using a large JSON file per mailbox containing all exported emails as input. For parallel processing, we used the GNU Parallel [40] tool. On behalf of the researchers, the university's IT staff executed the pipeline on the backup email server to make sure raw emails were not exposed to the researchers. Below, we describe the metadata we collected. In the cases where we applied pseudonymization, we provide a description of the pseudonymization procedure. Table IV in Appendix B gives an overview of both general and S/MIME and PGP specific information.

General Information. For each email we collected the local user account, message ID, the sender, and the list of receiver email addresses. If present, we also collected the lists of carbon copy and blind carbon copy addresses for outgoing emails. For pseudonymization, we hashed all of these values using a secret salt[10] and the SHA-256 hash function. We grouped email users into: Student, Staff, Faculty, NX Unknown[11] and External. For data protection reasons, we did not collect the exact send and receive dates as well as times of an email but only the corresponding week. If set, we collected the raw user agent string to identify email client software, operating system, and compute platform if possible. We grouped mailbox folders into: Inbox, Subfolder of Inbox, Outbox, Subfolder of Outbox, Junk, Trash, and Spam. For further cryptographic metadata analysis, we stored whether an email was signed and encrypted or contained Autocrypt headers. Below, we use the term *cryptographic emails* for all emails that contained cryptographic metadata.

Cryptographic Metadata. For all cryptographic emails, we collected attached cryptographic metadata for S/MIME, PGP and Autocrypt. Table IV gives an overview of the collected information and storage formats including pseudonymization.

For S/MIME, we collected the pseudonymized serial number of the leaf certificate, validity start and end date (granularity by week), the signing hash algorithm, the key size and key type (e.g., 2048 and RSA). We collected the key usage and extended key usage options (e.g., email signing,

[9]https://www.dovecot.org/

[10]The secret salt was only accessible to the IT staff and not to the researchers

[11]Used if the email subdomain did not exist anymore, and the original purpose was unclear when we performed our experiments.

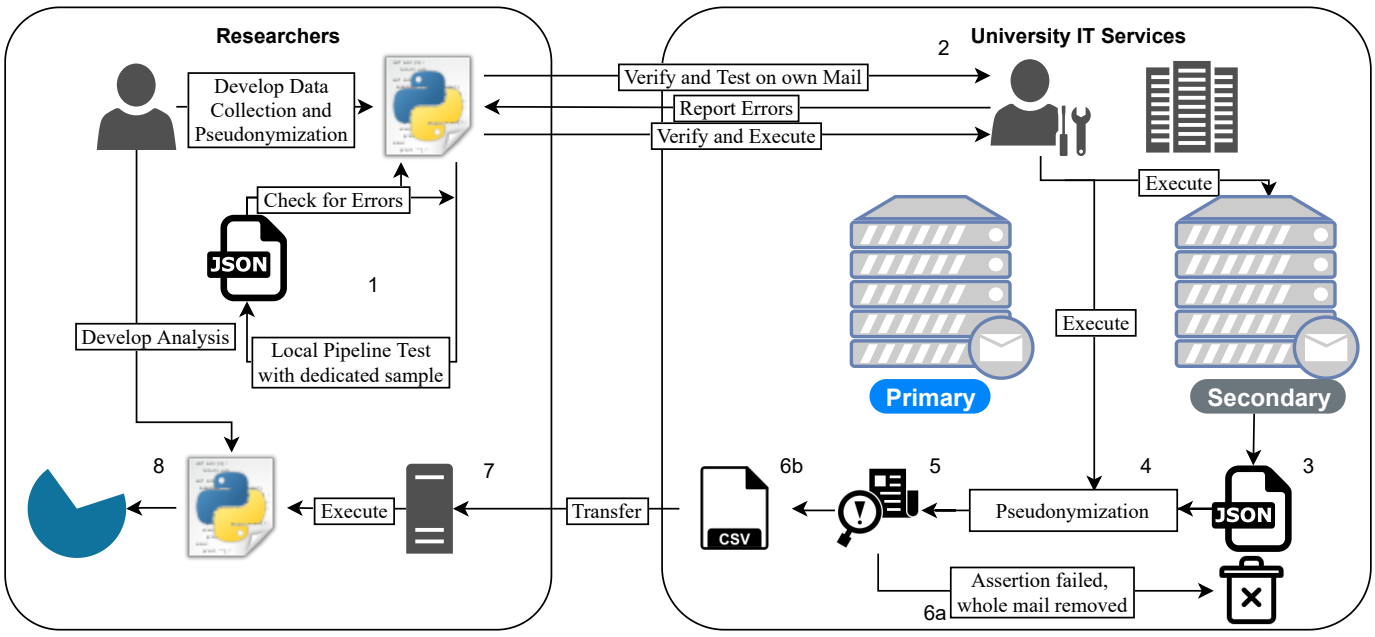

Fig. 1. Illustration of our data collection and evaluation pipeline.

certificate signing, code signing), and the number of valid email addresses for each certificate. Finally, we collected the complete certificate chain, including all metadata for all signing certificates. We did not pseudonymize the serial numbers for non-leaf certificates.

For PGP, we collected the key type (e.g., public key or a sub-key), the signature algorithm, and key length. For elliptic curve keys, we collected curve information as well as pseudonymized key IDs, and creation and expiration dates. For extended PGP keys, we collected update dates. If a key included subkeys we also stored the above data for the sub keys.

The pipeline dropped all data (like email subject, email content, non-key attachments) that we excluded from our analysis.

B. Data Cleaning

We included all 81,647,559 emails and 37,463 email user accounts at our university from January 1994 to July 2021 in the initial analysis (cf. Section V-A). However, we excluded some emails and email user accounts based on the following procedure:

Processing Errors. Parsing a dataset that spans millions of emails from over two decades that were generated and sent by many different email clients and versions poses a significant challenge. Hence, we were not able to successfully parse all emails. Our parser failed to parse 0.09% emails in the dataset. Due to privacy restrictions, we were not allowed to further investigate root causes of parser failures. However, Appendix A provides more details on S/MIME and PGP related parsing errors.

Inactive Email Accounts. Our initial data set included 37,463 total user accounts. However, we identified 18,302 inactive email accounts for which we did not find any sent emails. Of them, 17,928 email accounts received but did not send emails and 374 did neither receive nor send emails. Many students prefer to use their private email accounts instead of their automatically created university email accounts, leaving them inactive.

Invalid Dates. We excluded 307,680 (0.38% of our dataset) emails with obviously forged header dates, e.g., year <1994 or >2021 (after data collection) and emails for which the date parser failed.

C. Ethical Concerns and Data Privacy

To conduct the large scale measurement study on email data, our institutions, and specifically the institution where the data was collected and evaluated, did not require formal ethical review for this type of study. Therefore, we did not involve an ethics review committee. However, we followed our institutions' guidelines for *good scientific practice*, which includes ethical guidelines. Here, the institutions specifically place the burden of determination of whether research is ethical on the respective researchers. We intensively discussed within and outside our research team to determine possible concerns with this research project, and whether this project would be feasible. We concluded that in addition to following laws and our institutions' ethics requirements, we should also follow the de facto ethics standards of the S&P community. The data used in this study can be described as pseudonymized data derived from human subjects, as mentioned in the Call for Papers.

In addition to ethics, we made sure to address all legal aspects of our research to adhere to strict German privacy protection laws and the European General Data Protection Regulation (GDPR). Therefore, we involved the data pro-

tection officer and the works committee of the institution where the data was collected and evaluated, as required by the German data protection regulations. We developed the data collection plan jointly with the data protection officer, with the goal to protect users' privacy and adhere to the strict German data protection regulations and the regulations in the state of Lower Saxony. After more than a year of multiple discussions and hearings, we agreed on the presented data collection plan (cf. details in Section IV-A). After the involved authorities had rigorously assessed the legal situation based on the GDPR, German data protection laws, and state law of the involved authorities, we were allowed to analyze pseudonymized metadata of all users at our institution without requiring user consent. Additionally, our legal counsel decreed that the benefit of our research to society outweighed the risks to individuals. We concur with the assessment that answering our research questions is beneficial for future end-to-end encryption research, which ultimately benefits society, and that there was no harm done to any participants based on possibly re-identifiable metadata. Importantly, we will not publish the metadata we collected and only an encrypted copy will be stored at the university data center for ten years without access by researchers to follow good scientific practice. As is common in research, we only publish aggregate data, and no email accounts can be re-identified through the publication of this paper. As part of the joint development of our data collection plan, we decided to take the following measures to protect users' (metadata) privacy and adhere to the GDPR, German, and state laws:

- The involved researchers never had access to raw data. The data collection pipeline was executed by the university's IT staff who operate the email servers and have access to the backup data. They transferred the pseudonymized data to the researchers to a secure server.
- We reduced the amount of data to the absolute minimum we required to investigate our research questions.
- We used cryptographically secure hash functions with salts unavailable to the researchers to pseudonymize user data.
- At all times pseudonymized data (cf. Table IV in the appendix) was only stored on secured university servers.
- We did not and will not share any data other than the aggregate numbers in the paper with anyone outside the team of involved researchers.
- We assured the data protection officer and the works committee that we would not take any actions to de-identify users.

The above data protection precautions ensured the necessary compromise between user privacy and data utility in order to perform the analyses on which we report in Section V.

D. Limitations

Like every measurement study, our work comes with several limitations:

Data Set. Our dataset includes email data from January 1994 to July 2021 covering 81,647,559 emails of 37,463 users from a large Germany University with more than 30,000 students and 5,000 employees. However, the dataset might not be complete. We could not include emails that users deleted from their accounts; similarly, we did not include emails sent to deleted accounts. Some research groups and departments have their own email servers; their emails were also not included in our dataset. Therefore, the dataset should not be assumed to include all emails that were sent or received from January 1994 to July 2021 at our university. Additionally, we do not assume that our dataset is representative for all email data in Germany or globally. Instead, we think our data set might overreport the use of email encryption since most email users in the dataset are highly educated and our university offers free S/MIME certificates to all email users.

However, we think our dataset is one of the largest and most valuable for the type of analysis we perform and our results are a valuable contribution to the security research community.

Data Analysis. We could not analyse all cryptographic details, e.g., we could not verify digital signatures since we were not allowed to parse the body content of emails and we could not extract details for certificates or signatures that were used in encrypted emails, since this data is also encrypted.

We tested our pipeline thoroughly but still missed some edge cases that inevitably arise in mail software that evolves over a long time span. While our pipeline was able to process 99.91% of all emails, processing failed for the remaining 0.09%. Errors during S/MIME and PGP parsing were logged separately. We encountered 1,199 S/MIME and 23,168 PGP emails where parsing failed (cf. Appendix A for more details). We deemed this margin of error tolerable compared to the high organizational costs of refining and repeating the entire process once more.

Distinction Between Send and Receive. The dataset we evaluated did not contain information whether an email was sent or received. To still group emails into sent and received emails, one approach would be to group emails based on the folder they were stored in. However, this would introduce challenges such as email clients using different names for sent folders or users using their own folder names. Therefore, we decided to identify emails based on multiple parameters. Sent emails were not allowed to contain a `return_path` header, since it is added by outgoing mail servers and emails were only allowed to go through at most one mail relay[12]. Most incoming emails have more than five email relays.

Mail Client Behaviour. The macOS and iOS clients AppleMail, iPhoneMail and iPadMail generally identify themselves using the *X-Mailer* header in mails, but the copy placed in the sent folder by these clients does not contain this header. As a result, these clients could only be correctly detected on received emails, their sent mails are included in the "No User

[12]One mail relay header is added when the webmail client of the university is used for sending mails. Regular mail clients do not add this header.

Agent" group. The ticket system used by the university data center automatically deletes mails in its inbox and does not place a copy in the sent folder and as a result only the answered tickets are available in our dataset.

E. Replication Package

To improve the replicability of our work, we provide a replication package including the following material: (a) the complete processing pipeline consisting of multiple Python scripts to process and pseudonymize emails from Dovecot mail servers, (b) the analysis scripts to replicate our results on different datasets, and (c) the agreement with our data protection officer. Due to the sensitive nature of our measurement study, we cannot make raw email data available. We hope this replication package helps future studies to better compare and position themselves to our work; and hope others replicate our work on different email datasets to improve our community's understanding of the use of email encryption. The replication package is available on our website at [39].

V. Results

We provide a detailed analysis of the email corpus we collected and the adoption of S/MIME and PGP from 1994 to July 2021 below.

A. Dataset Summary

In total, we analyzed metadata for 81,647,559 emails from 37,089 email accounts. Overall, the university's email users exchanged 40,540,140 (49.67%) emails internally.

Figure 2 illustrates the use of email at our university in the past 27 years. While we found only 350 emails in 1994, we detected an almost exponential growth and found 17,190,472 emails in 2020. This development reflects the enormous relevance of email as a communication tool and is in line with reports on the global use of email[13].

Use of Email Encryption and Signatures. We found 2,334,042 (2.86%) emails that were either encrypted or signed using S/MIME or PGP. We identified a huge difference between the use of email encryption and signatures.

46,973 (0.06%) emails were encrypted. 26,105 (55.57%) emails were encrypted using S/MIME and 20,868 (44.43%) of them were encrypted using PGP. In contrast, 2,287,922 (2.8%) emails were signed. 2,040,794 (89.2%) of them were signed using S/MIME and only 247,128 (10.8%) were signed using PGP.

Figure 2 illustrates the use of S/MIME and PGP between 1994 and 2020. We found the first S/MIME signed email in 1998 and the first S/MIME encrypted email in 1999. The first PGP signed email appeared in 1994 and the first PGP encrypted email in 1997.

> Key Insights: Dataset.
> - We saw an exponential growth of the use of email between 1994 and 2020.
> - Only 0.06% of emails were encrypted.

[13] cf. https://www.emailisnotdead.com/

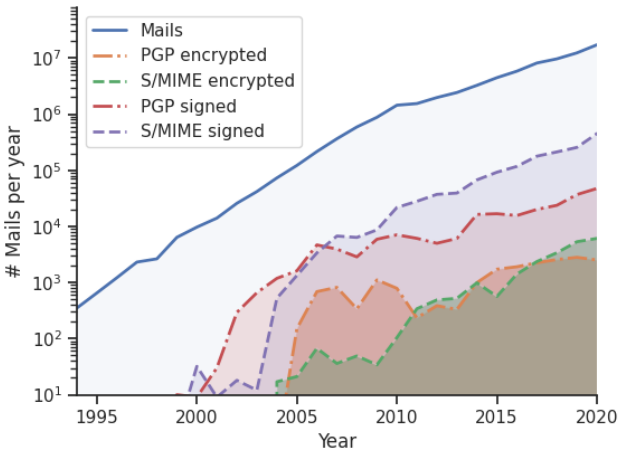

Fig. 2. Rise of email, S/MIME and PGP over time at our university.

> - 2.8% of emails were signed.
> - S/MIME was more widely used than PGP.

B. S/MIME Certificates and PGP Keys

Below, we give an overview of the S/MIME certificates and PGP keys we found. This includes certificates and keys from internal and external senders. Overall, we were able to collect 9,765 S/MIME certificates, 3,741 primary PGP keys and 3,840 sub keys (cf. Table II for details).

S/MIME Certificates. All but one certificate that used an elliptic curve encryption algorithm supported the RSA encryption algorithm.

2048 bits was the most widely used RSA key size (91.58%); 5.54% of the RSA keys had 4096 bits. In 237 cases we saw 1024 bits RSA keys; 6 RSA keys had 512 bits. While the last 512 bits RSA key we found was created in 2010, we saw 2 1024 bits RSA keys created in 2020.

7,472 (76.52%) certificates supported the SHA-256 signature algorithm. However, we also found outdated signature algorithms including SHA-1 (2,028; 20.77%) or MD5 (148; 1.52%). Surprisingly, we found 11 certificates issued in 2020 using SHA-1. The last certificate using MD5 was generated in 2017.

5,194 (53.19%) of all certificates expired in 2020 or earlier. The mean validity period for S/MIME certificates was 3.13 years (sd= 2.70) ranging overall from a minimum of 4.00 weeks to a maximum of 99.99 years. 6,953 (71.20%) certificates were created between 2015 and 2020 with a peak of 1,654 certificates (16.94% of all S/MIME certificates) in 2019.

Overall, we found 671 different issuer names. However, 1,150 (11.78%) certificates had no issuer. The most prominent issuer was the DFN issuing 3,209 (32.86%) certificates. 622 (6.37%) were issued by our university. Another German university issued 563 (5.77%) of all S/MIME certificates we found. In 332 (3.40%) cases, a distinct issuer only signed a single certificate. 89 (0.91%) issuers signed only two certificates. 137 of them (32.54%) had no root CA. In total, we

	S/MIME				PGP				Total	
	Sent		Received		Sent		Received		Sent	Received
	Signed	Encrypted	Signed	Encrypted	Signed	Encrypted	Signed	Encrypted		
Total	356,330	9,358	1,684,464	16,747	69,950	8,197	176,983	12,633	16,660,280	64,952,315
Client										
Thunderbird	231,475	6,311	508,423	7,103	46,407	5,775	85,904	6,825	8,206,215	11,875,969
Outlook	78,736	1,675	258,013	3,728	325	22	5,528	52	3,102,760	7,785,073
Ticketsystem	0	0	311,743	1	0	0	223	0	4	1,259,208
AppleMail	?[1]	?[1]	58,752	2,157	?[1]	?[1]	20,722	1,408	?[1]	2,268,757
Evolution	11,479	190	16,366	341	232	48	1,387	74	26,224	70,482
Mutt/NeoMutt	0	29	50	31	564	1	11,506	1	6,066	141,263
Outlook-Express	7	0	7,355	31	0	0	2,452	16	29,065	447,775
Claws Mail	0	0	145	0	1,654	207	2,918	161	2,573	21,640
iPhone/iPad-Mail	?[1]	?[1]	5,275	104	?[1]	?[1]	7	8	?[1]	723,482
MailMate	2,361	8	3,526	3	63	2	131	3	5,296	8,364
Other	1,270	3	13,397	128	2,804	221	8,658	2,167	2,572,631	9,715,649
No Useragent	31,002	1,142	501,419	3,120	17,901	1,921	37,547	1,918	2,708,758	30,634,653
Operating System[2]										
Windows	167,350	2,183	326,747	2,555	38,751	2,531	66,786	2,589	7,150,676	10,679,520
Linux	60,827	3,985	164,926	4,379	8,797	3,372	23,069	4,191	799,179	1,457,578
Mac	5,909	170	88,144	2,397	2,126	172	24,849	1,841	340,599	3,057,579
iOS	1	0	5,294	104	0	0	7	8	695	726,209
Android	81	1	66	0	34	101	167	110	84,744	120,115
Webmail	135	3	831	8	10	2	322	2	2,378,940	2,425,629
Unknown	91,025	1,874	597,037	4,184	2,331	98	24,236	1,974	3,196,689	15,851,032
No Useragent	31,002	1,142	501,419	3,120	17,901	1,921	37,547	1,918	2,708,758	30,634,653
Usergroup										
Scientific	237,579	3,992	808,539	8,376	66,483	6,708	115,388	7,796	11,776,198	39,504,464
Staff	106,434	5,318	684,791	8,155	2,122	1,134	49,111	3,864	3,265,602	17,560,401
NX Internal	9,903	20	121,243	129	77	39	7,153	436	970,160	4,960,051
Student	1,700	28	67,813	71	1,242	311	4,903	536	393,685	2,703,240
External	714	0	2,078	16	26	5	428	1	72,256	224,159
Affected emails	>= 5%	> 4%	> 3.5%	> 3%	> 2.5%	> 2%	> 1.5%	> 1%	> 0.5%	<= 0.5%

[1] AppleMail and the iOS mail client miss the *X-Mailer* header in the sent folder. Hence, we identify their client as "No Useragent".
[2] Not all clients store the operating system as part of the Useragent/X-Mailer field. If not available, we identify them as "Unknown".

TABLE I
TOP 10 CLIENTS OF S/MIME AND PGP USERS AT OUR UNIVERSITY.

identified 495 root CAs. 274 of them were only the root of one certificate. The Deutsche Telekom was the root CA of 6,342 (64.95%) certificates. In contrast, 1,150 (11.78%) were self-signed and not linked to a CA.

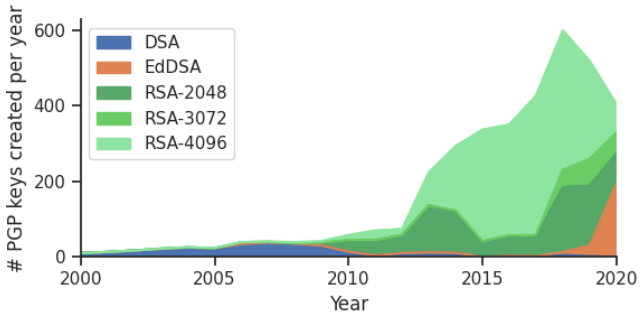

Fig. 3. Distribution of primary PGP keys over time.

PGP Keys. PGP keys can be split into primary and sub keys. Each primary key can have multiple sub keys. Below, we focus on primary keys. We found 3,741 primary and 3,840 sub keys. Primary keys had on average 1.03 sub keys with a maximum of 10.

Most PGP keys used RSA (3,169; 84.71%) and 2048 bits (928; 24.81% keys) or 4096 bits (2,015; 53.86% keys) keys, followed by DSA (323; 8.63%) with mostly short key sizes <= 1024 bit (290; 7.75%) for signatures and Elgamal sub keys for encryption.

PGP key expiration dates can be set and extended by the user themselves. We saw this in 219 (5.85%) cases with a maximum of 4 extensions. 1,234 (32.99%) keys had no expiration date, while 1,344 (35.93%) keys were expired. The average validity period for PGP keys was 4.67 years (sd 3.39) ranging from zero weeks to 85.82 years.

The majority (276; 85.45%) of DSA keys were created before or in 2010 and used 1024 bit keys. 278 (86.07%) had no expiration date set. The last two DSA keys were created in 2019 with 3072 bit keys.

928 (24.81%) of the RSA keys used 2048, 2,015 (53.86%) 4096 bits. Only 6 (0.16%) had 1024 bit keys. 2048 bit was the most widely used key size (60-70% per year) between 2010 and 2013. 4096 bit was most dominantly used from 2014 and peaking in 2017 with 86.18%. Interestingly, in 2018 the picture changes again and 2048 and 3072 bit keys appeared more frequently again with 2048 bits taking back the lead in 2020, although the amount of RSA keys is starting to decline in favor of EdDSA. We attribute this phenomenon to the release of Autocrypt in 2018 which uses RSA 2048 bit keys by default to avoid the 10kB header limit. [6].

			# S/MIME	# PGP	
				Primary	Sub
			9,765	3,741	3,840
Cryptographic Algorithms[1]	RSA	512	6	-	-
		1024	237	6	6
		2048	8,942	928	950
		3072	32	208	226
		4096	541	2,015	2,078
		8192	5	5	2
	DSA	768	-	1	-
		1024	-	290	1
		2048	-	17	2
		3072	-	15	2
	ElGamal	768	-	-	1
		1024	-	-	61
		2048	-	-	186
		3072	-	-	13
		4096	-	-	53
	EC	521	1	-	-
	ECDH	256	-	-	247
	EdDSA	256	-	249	2
Validity Period	mean	(years)	3.13	4.67	4.62
	std. dev.	(years)	2.70	3.39	3.68
	median	(years)	2.99	4.99	4.99
	min	**(weeks)**	4.00	0.00	0.00
	max	(years)	99.99	85.82	85.82
	unlimited	-	-	1,234	1,240

[1] Key sizes in bits

TABLE II
DETAILS FOR S/MIME CERTIFICATES AND PGP KEYS IN OUR DATASET

We found the first four EdDSA keys in 2017, followed by nine in 2018, 31 in 2019 and 193 in 2020, almost reaching the number of RSA keys (207). 53 (21.29%) of the EdDSA keys did not have an expiration date set.

Until 2000, all 33 PGP mails were hashed with MD5. Between 2004 and 2015, SHA1 was most widely used in 62,273 (80.96%) PGP mails.

We found the first use of SHA2-256 in 2006 with a steady increase until 2016 when it became the most widely used algorithm with a total of 125,973 mails (68.54% since 2016; 50.22% overall). SHA2-512 first appeared in 2008, has constantly been growing since, and has been the second most used hash algorithm since 2016 (44,859; 25.94% since 2016; 20.26% overall). The remaining hash algorithms SHA-224 (7; <0.01%) and RIPE-MD160 (132; 0.05%) were almost non-existent in our dataset.

Key Insights: S/MIME and PGP.
- RSA is the most widely used key algorithm.
- A key size of 2048 bits was used most often with S/MIME.
- PGP keys used 4096 bits most often, although newer PGP keys used less secure 2048 bits.
- About one third of the PGP keys did not have an expiration date set.
- The *Deutsche Telekom* was root CA for 64.95% of all S/MIME certificates.

C. User Interaction with S/MIME and PGP

Overall, we identified 37,463 unique user accounts including 374 accounts with no emails at all and 18,302 users who never sent an email. Below, we focus our analysis on the remaining 19,161 active email users who have sent at least one email. Active users sent 688.93 emails (ranging from 1 to 182,379 emails) and received 3,249.81 (ranging from 0 to 2,881,904 emails) on average.

Based on the subdomain of user accounts' postboxes we assigned user accounts to the groups *scientific*, *staff*, *students*, *external* and *NX internal*. We assigned administrative staff of a research group to *scientific*, assigned subdomains that were no longer available when we collected and analyzed the data to *NX internal* and *external* to accounts that are hosted by the university but belong to external research projects. 9,053 (24.41%) of our users were in the scientific user group, 2,169 (5.85%) in staff and 19,002 (51.23%) were students (cf. Table I for a distribution of emails for all user groups).

Below, we focus on user behavior. We aim to understand real world implications of the common usability challenges (cf. [5], [20], [21], [27], [45]) with a focus on:

- General use of S/MIME and PGP.
- Use of S/MIME and PGP with multiple clients.
- Distributing S/MIME certificates and PGP keys.
- Long-term S/MIME certificate and PGP key management.
- Leakage of private keys.

General Use of S/MIME and PGP. 94.54% of the active users never used S/MIME or PGP for sending email. In contrast, 62.59% of our active users received at least one email signed or encrypted using PGP or S/MIME.

375 users (1.96% of the active users) sent at least one encrypted email. While 167 users used PGP exclusively and 159 used S/MIME exclusively, 49 used both. 1,047 users (5.46% of all active users) signed at least one email. 446 users PGP exclusively, 455 users used S/MIME exclusively, and 146 users used both.

S/MIME users signed 33.58% of their emails on average after using S/MIME for the first time but encrypted only 1.26% of their emails on average. PGP users signed 4.90% of their emails on average but encrypted only 0.96% of their emails.

Breaking down S/MIME and PGP use into user groups, we observed that staff users used S/MIME to sign their emails more often than scientific users (3.26% vs 2.02%) and eight times more often than students (0.43%) on average. In contrast, the amount of PGP signed emails is smaller across all user groups, with the scientific user group being at the top with 0.56%, the staff users with 0.06%, and students with 0.32% on average.

For more details including the use of encryption, please refer to Table I.

Use of Multiple Email Clients. To identify email clients, we relied on user agent metadata in emails. We were able to detect email clients for 48,269,184 (59.14%) emails. Hence, we could detect the email client for 13,951,522 (83.74%) emails sent by our users, for 14,492 (82.55%) encrypted and 377,377 (88.53%) signed emails. However, 33,343,411 (40.86%) emails did not contain user agent information.

The most common client used for signing and encryption was Thunderbird in combination with Enigmail for PGP support[14]. It was used to send 289,968 (65.35%) S/MIME and PGP emails in our dataset. 80,758 (18.20%) of the S/MIME and PGP emails were sent using Outlook. Table I provides more details on the top 10 email clients.

Using multiple email clients posed additional burden on the widespread use of S/MIME and PGP since users have to synchronize their keys and certificates across all clients. For 8,828 (46.07%) out of 19,161 users, we could detect multi-client use. Table III provides an overview of multi-client use across all users and S/MIME and PGP users. Figures 5 and 6 compare how S/MIME and PGP users employ signatures and encryption between single- and multi-client users. Our data implies that single-client users sign emails more frequently than multi-client users. For S/MIME users, the number of signed emails decreases from median 62.25% for single-client users to median 1.58% for dual-client and to 1.03% for triple-client users. However, for power users with more than three clients the number of signed emails increases to median 17.48%. The majority (74.90%) of multi-client users tend to use one of their clients for email signatures exclusively and not employ signatures with other clients, while only 25.10% of the multi-client users used email signatures across different clients. In contrast, single-client PGP users only signed 0.36% of their emails on average. However, we did not find considerable differences between single and multi-client PGP users (cf. Table III for more details).

S/MIME Rendezvous. In addition to multiple clients, the distribution of S/MIME certificates and PGP keys between users poses another significant challenge for the adoption of email encryption. We focus on S/MIME certificate distribution, since S/MIME clients automatically attach certificates including public keys to signed emails. However, plain PGP clients[15] do not automatically attach their public keys to signed emails but require users to manually attach public keys or look them up on PGP key servers.

Signed S/MIME emails were the most frequent cryptographic emails in our dataset. Therefore, we were able to investigate users' interactions with S/MIME certificates they received from their communication partners. Without additional effort for key exchange, S/MIME users who received S/MIME certificates from others could encrypt future emails to those senders. Below, we report on such behavior of users in our dataset. 601 users (3.14% of all internal active users) sent at least one signed email. Overall, we identified 374 rendezvous where both sender and recipient exchanged public S/MIME keys due to sending signed emails to each other. Most emails (64.08%) between S/MIME rendezvous partners were signed. However, only 3.36% of all emails between rendezvous partners were encrypted on average. Once one rendezvous partner had sent a first encrypted email, 13.95% of all following emails were encrypted on average. Figure 4 illustrates the distribution of signed and encrypted emails between S/MIME rendezvous partners.

[14]Native PGP support was added to Thunderbird 78 (June 2020)

[15]These clients do not have Autocrypt support

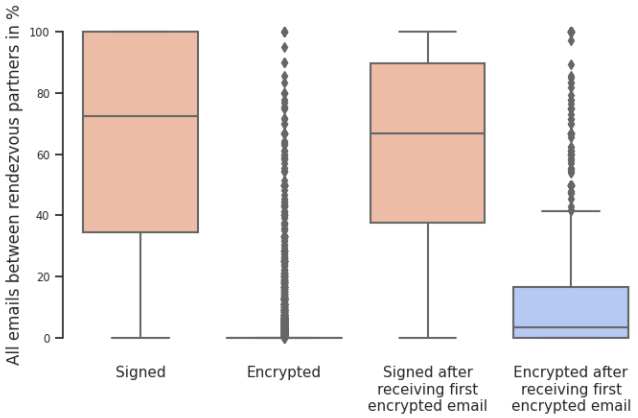

Fig. 4. Distribution of signed and encrypted emails between S/MIME rendezvous partners. While they signed most of their emails, they encrypted only few. However, receiving an encrypted email had a positive impact on encrypting future emails between rendezvous partners.

Long-term S/MIME and PGP Key Management. Below, we report on the long term use of S/MIME certificates in our dataset. In particular, we focus on the replacement of outdated certificates. Overall, we identified 680 university email addresses with at least one S/MIME certificate which was actively used. 496 (72.94%) of them were valid until December 2020. 203 (29.85%) of these email addresses had two or more certificates associated. On average, they used 2.86 certificates – one used up to 27 certificates.

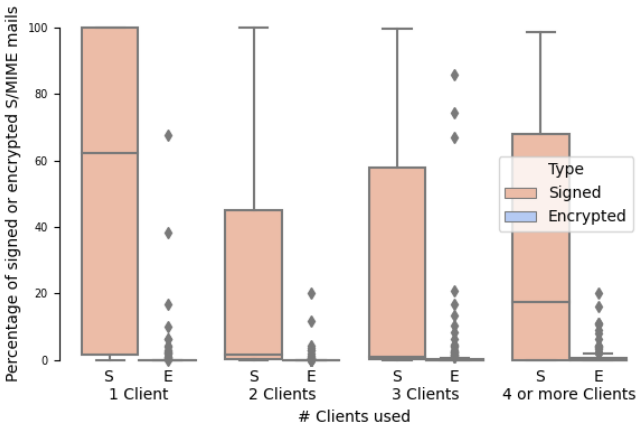

Fig. 5. Distribution of S/MIME signatures and encryption for users with one or multiple clients.

Since certificates expire and need to be replaced with new ones, users need to create, set up, and distribute them ideally very close to the expiration date of the old certificate. Overall, we found 364 certificate rollovers. In 271 cases, the new certificates had a longer expiration period than the old certificates. 229 of the certificate rollover events we detected occurred in time before the old certificates expired. On average, they happened 55.05 weeks before the old ones expired. Figure 7 shows that the majority of users often create

	S/MIME								PGP							
	Emails signed[1]				Emails encrypted[1]				Emails signed[1]				Emails encrypted[1]			
clients[2]	1	2	3	4+	1	2	3	4+	1	2	3	4+	1	2	3	4+
mean	53.55	23.50	28.17	33.22	1.39	0.33	2.57	1.02	14.08	6.01	2.08	4.39	2.53	0.97	0.72	0.75
std. dev.	44.99	35.79	36.09	35.46	7.17	1.88	11.38	2.66	27.05	18.14	8.84	13.90	12.91	2.91	2.65	1.81
25%	1.59	0.14	0.13	0.10	0.00	0.00	0.00	0.00	0.08	0.04	0.03	0.03	0.00	0.00	0.00	0.00
50%	62.25	1.58	1.03	17.48	0.00	0.00	0.00	0.06	0.36	0.13	0.10	0.09	0.00	0.00	0.00	0.00
75%	100.00	45.03	57.92	67.91	0.00	0.00	0.24	0.79	10.00	0.57	0.32	0.42	0.00	0.00	0.00	0.48
max	100.00	100.00	99.48	98.61	67.59	20.00	85.85	20.17	100.00	93.75	72.61	91.23	100.00	20.53	22.58	13.52

[1] Mean, std. dev, percentiles and max values in %
[2] Number of clients used by single users.

TABLE III
DISTRIBUTION OF SIGNED AND ENCRYPTED EMAILS FOR MULTI-CLIENT USERS.

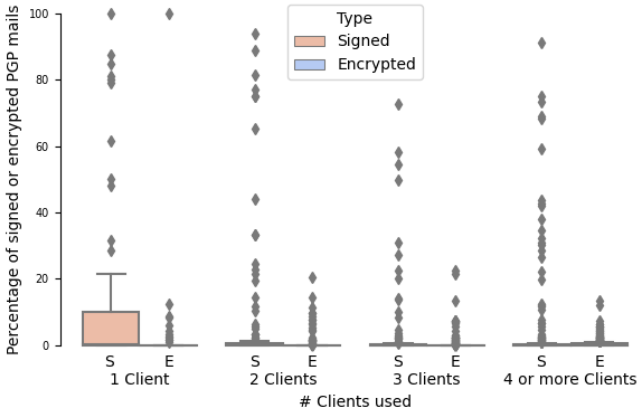

Fig. 6. Distribution of PGP signatures and encryption for users with one or multiple clients.

a new certificate shortly before the previous certificate expires. However, we also detected 42 certificate rollover events that occurred after the old certificates expired. S/MIME could neither send signed nor receive encrypted emails in this time period. On average, the late rollover events occurred 70.64 weeks after the expiration dates.

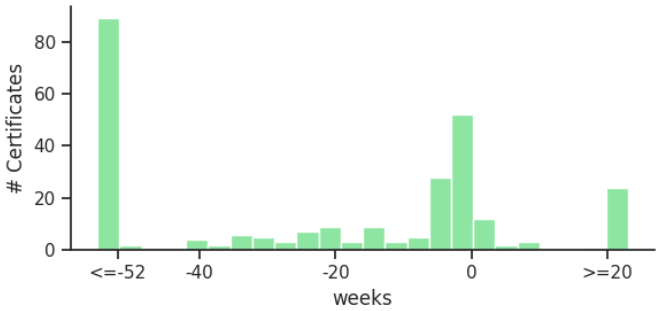

Fig. 7. Distribution of the time around certificate expiration for certificate renewals in weeks. Most certificates were renewed around one year before they expired due to an expired root CA certificate. The second most certificates were renewed one week before they expired.

Leakage of Private Keys. Overall, we encountered three instances of private PGP keys (and their private sub keys) being sent via email in 2015, 2017, and 2018. All three keys were sent by the users to themselves. One of those was a freshly created PGP key (i.e., less than one week old). This is not necessarily a security issue as long as the passphrase to protect the key is adequate.

Key Insights: S/MIME and PGP users.

- More than 94% of all active users never used S/MIME or PGP.
- S/MIME users signed six times more of their emails than PGP users on average.
- Using two to three different clients decreased the likelihood of signing emails by 51.76%.
- On average, less than 3% of all emails between users who had exchanged S/MIME certificates previously were encrypted.
- Leakage of private keys via email does not seem to pose an issue.

VI. DISCUSSION

In this section, we discuss the implications of our results and recommendations for the future development of email encryption in five key areas: Limited use of email encryption, use of insecure keys, impact of certain events, challenges of using multiple clients, and lack of opportunistic encryption.

Limited Use of Email Encryption. As illustrated in Section V-A, we observed that only a very small fraction of emails in our dataset was encrypted (0.06%) or signed (2.8%). While we saw an exponential growth of the use of email overall (cf. Figure 2), the fraction of encrypted emails remained consistently small. Our results also imply that S/MIME was more widely used than PGP for both signing and encrypting emails. Although these findings do not come unexpected in the light of previous work [22], [26], [31], [32], [35], we think they can serve as ground truth to confirm previous user study results and have value for future development. The small fraction of email encryption and the fact that the use of email encryption did not grow with the overall use of emails suggest that both S/MIME and PGP are niche tools that are mostly used by a small number of security-aware users. However, the difference between the use of S/MIME and PGP seems to give grounds for optimism. Our institution provides an S/MIME infrastructure and encourages its use to improve security and privacy without advertising it aggressively. These findings are in line with results of previous works [1], [2] which find that many users do not use end-to-end encryption for security reasons but adopt it along with other features. Therefore, future development should look into adding more value to the use of email encryption on top of security and privacy. For example, more widely accepting digitally signed

emails for administrative processes could increase the value of email encryption for a broader set of users.

Use of Insecure Keys. For both S/MIME and PGP, we found only small numbers of insecure or outdated keys (cf. Table II). The majority of keys used sufficient key sizes (>= 2048 bits) for RSA and Elgamal and >= 1024 bits for the DSA algorithm. However, we found a significant number of keys that had no expiration date set or were used after their expiration date. In particular, most PGP keys were affected (32.99% had no expiration date and 35.93% were expired). While the widespread adoption of secure algorithms and sufficient key sizes is positive, the prevalence of key expiration issues is a matter of concern. These findings suggest that initial key generation by email encryption clients and plugins works well and supports email users in setting up sufficiently strong encryption keys. However, they also illustrate that tool support to prevent the use of insecure and expired keys has limits. Therefore, future development should look into better prevention mechanisms that nudge users into only using secure and still valid keys.

Impact of Certain Events. Our longitudinal field data allows us to shed light on the impact of specific events in time on the use of email encryption. In the following, we focus on two events: The Snowden revelations in 2013 and the COVID-19 pandemic in 2020. While this list is not exhaustive, both events had significant impact on the security community.

In 2013, Edward Snowden leaked the mass-surveillance program of the NSA and other security agencies [23], [24]. Users of modern digital communication tools were made aware of the significance of end-to-end encryption for information security. Compared to 2012 (0.035%), the use of email encryption doubled to 0.07% in the following years. While the impact of the revelations on absolute numbers is limited, the use of email encryption still significantly increased. This incident illustrates that awareness campaigns may positively affect the adoption of end-to-end encryption which is in line with an unprecedented growth of Signal users after a WhatsApp controversy in January 2021[16].

During the COVID-19 pandemic, we found that more emails were exchanged than before. This can be explained by the fact that most of the institution's staff worked from home and many administrative processes that had been paper-based before the pandemic were digitalized in 2020. Hence, for all emails we noted an increased amount of emails sent (39% compared to 2019). However, the number of S/MIME-signed emails also increased by almost 76% from 255,104 to 449,646 compared to 2019. In contrast, the amount of encrypted emails slightly dropped from 0.061% in 2019 to 0.050% and 0.051% in 2020 and 2021, respectively. Users might have had problems setting up and using email encryption from home and accessing the university's technical support during the pandemic. Similarly, about one-third fewer certificates were issued to users in 2020 compared to the previous year (cf. Figure 10).

[16]cf. https://www.reuters.com/article/us-signal-users-idUSKBN29I27U

Key Management Challenges. We identified multiple key management challenges. While we found limited use of email encryption in general, the use of multiple clients posed additional challenges for S/MIME users in our dataset (cf. Section V-C) – single client S/MIME users signed more than 60% while users with two clients only signed 1.58% of their emails, which might be caused by a lack of tool support for transferring S/MIME setups between clients. Similarly, we identified challenges for long-term key and certificate management. 42 out of 364 certificate rollovers occurred months after the certificates' expiration dates, which made it impossible to produce valid signatures in the meantime. Better tool support for easier key rollovers could contribute to earlier certificate updates for a broader set of users. However, we also identified a positive aspect of key management. Only three emails had private keys attached, all of which were sent to the users' own mailboxes. Hence, in no case was private key material leaked to unauthorized users. Overall, while our data implies that key management needs to be improved, private key leakage through email attachments was not an issue for our users.

Missed Opportunities. In addition to the key management issues above, we identified the initial distribution and use of keys and certificates to be challenging for users. In particular, using already exchanged public keys between users for further email encryption was a surprising issue we found. While we identified mutual public S/MIME key exchanges between 374 users, they only encrypted 13.95% of their future emails. Overall, 318,214 emails in our dataset could have been encrypted between these users without any additional key exchange. Here, our data illustrates that an automated mechanism for more efficient opportunistic encryption similar to Autocrypt could potentially to help to increase the number of encrypted emails.

VII. Conclusion

In this work, we presented the first analysis of a large corpus of longitudinal email data for thousands of users at a large German university. We were able to confirm common beliefs and results from previous work in the security community: Only few users used email encryption to secure only a small fraction of their emails. We identified key management to be challenging in particular in the context of multiple clients, key rollovers and key exchange. Based on our evaluation, we make suggestions for improving email encryption adoption. Overall, we hope our investigation provides a data driven motivation for future work to improve both the security and usability of email encryption solutions.

Acknowledgment

The authors would like to thank the staff at the Leibniz University IT Services at Leibniz University Hannover.

REFERENCES

[1] R. Abu-Salma, E. M. Redmiles, B. Ur, and M. Wei, "Exploring User Mental Models of End-to-End Encrypted Communication Tools," in *8th USENIX Workshop on Free and Open Communications on the Internet (FOCI 18)*. Baltimore, MD: USENIX Association, 2018.

[2] R. Abu-Salma, M. A. Sasse, J. Bonneau, A. Danilova, A. Naiakshina, and M. Smith, "Obstacles to the Adoption of Secure Communication Tools," in *2017 IEEE Symposium on Security and Privacy, SP 2017, San Jose, CA, USA, May 22-26, 2017*, 2017, pp. 137–153.

[3] O. Akgul, W. Bai, S. Das, and M. L. Mazurek, "Evaluating In-Workflow Messages for Improving Mental Models of End-to-End Encryption," in *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Aug. 2021, pp. 447–464. [Online]. Available: https://www.usenix.org/conference/usenixsecurity21/presentation/akgul

[4] T. Alrashed, A. H. Awadallah, and S. Dumais, "The Lifetime of Email Messages: A Large-Scale Analysis of Email Revisitation," in *Proc. 2018 Conference on Human Information Interaction & Retrieval*, ser. CHIIR '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 120–129.

[5] E. Atwater, C. Bocovich, U. Hengartner, E. Lank, and I. Goldberg, "Leading Johnny to Water: Designing for Usability and Trust," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*. Ottawa: USENIX Association, Jul. 2015, pp. 69–88.

[6] Autocrypt team, "Autocrypt Level 1: Enabling encryption, avoiding annoyances," https://autocrypt.org/level1.html.

[7] N. Avigdor-Elgrabli, R. Gelbhart, I. Grabovitch-Zuyev, and A. Raviv, "More than Threads: Identifying Related Email Messages," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, ser. CIKM '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1711–1714.

[8] W. Bai, D. Kim, M. Namara, Y. Qian, P. G. Kelley, and M. L. Mazurek, "Balancing Security and Usability in Encrypted Email," *IEEE Internet Computing*, vol. 21, no. 3, pp. 30–38, 2017.

[9] W. Bai, M. Namara, Y. Qian, P. G. Kelley, M. L. Mazurek, and D. Kim, "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)*. Denver, CO: USENIX Association, Jun. 2016, pp. 113–130.

[10] J. Callas, L. Donnerhacke, H. Finney, D. Shaw, and R. Thayer, "OpenPGP message format (RFC 4880)," 2007.

[11] J. Clark, P. C. van Oorschot, S. Ruoti, K. E. Seamons, and D. Zappala, "Securing Email," *CoRR*, 2018. [Online]. Available: http://arxiv.org/abs/1804.07706

[12] D. Davis, "Compliance Defects in Public Key Cryptography," in *Proceedings of the 6th USENIX Security Symposium, San Jose, CA, USA, July 22-25, 1996*. USENIX Association, 1996.

[13] Z. Durumeric, D. Adrian, A. Mirian, J. Kasten, E. Bursztein, N. Lidzborski, K. Thomas, V. Eranti, M. Bailey, and J. A. Halderman, "Neither Snow Nor Rain Nor MITM...: An Empirical Analysis of Email Delivery Security," in *Proceedings of the 2015 Internet Measurement Conference*, ser. IMC '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 27–39.

[14] S. Dusse, P. Hoffman, B. Ramsdell, L. Lundblade, and L. Repka, "RFC2311: S/MIME Version 2 Message Specification," 1998.

[15] M. Elkins, D. D. Torto, R. Levien, and T. Roessler, "RFC3156: MIME Security with OpenPGP," 2001.

[16] S. Fahl, M. Harbach, T. Muders, M. Smith, and U. Sander, "Helping Johnny 2.0 to Encrypt His Facebook Conversations," in *Proc. 8th Symposium on Usable Privacy and Security (SOUPS'12)*. ACM, 2012.

[17] D. Fisher, A. J. Brush, E. Gleave, and M. A. Smith, "Revisiting Whittaker & Sidner's "Email Overload" Ten Years Later," in *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work*, ser. CSCW '06. New York, NY, USA: Association for Computing Machinery, 2006, p. 309–312.

[18] I. D. Foster, J. Larson, M. Masich, A. C. Snoeren, S. Savage, and K. Levchenko, "Security by Any Other Name: On the Effectiveness of Provider Based Email Security," in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 450–464.

[19] S. L. Garfinkel, D. Margrave, J. I. Schiller, E. Nordlander, and R. C. Miller, "How to Make Secure Email Easier to Use," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '05. New York, NY, USA: Association for Computing Machinery, 2005, p. 701–710.

[20] S. L. Garfinkel and R. C. Miller, "Johnny 2: a user test of key continuity management with S/MIME and Outlook Express," in *Proc. 1st Symposium on Usable Privacy and Security (SOUPS'05)*. ACM, 2005.

[21] S. L. Garfinkel, J. I. Schiller, E. Nordlander, D. Margrave, and R. C. Miller, "Views, Reactions and Impact of Digitally-Signed Mail in e-Commerce," in *Financial Cryptography and Data Security*, A. S. Patrick and M. Yung, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 188–202.

[22] S. Gaw, E. W. Felten, and P. Fernandez-Kelly, "Secrecy, Flagging, and Paranoia: Adoption Criteria in Encrypted Email," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '06. New York, NY, USA: Association for Computing Machinery, 2006, p. 591–600.

[23] G. Greenwald and E. MacAskill, "NSA Prism program taps into user data of Apple, Google and others," *The Guardian*, vol. 7, no. 6, pp. 1–43, 2013.

[24] G. Greenwald, E. MacAskill, and L. Poitras, "Edward Snowden: the whistleblower behind the NSA surveillance revelations," *The Guardian*, vol. 9, no. 6, p. 2, 2013.

[25] A. Lerner, E. Zeng, and F. Roesner, "Confidante: Usable Encrypted Email: A Case Study with Lawyers and Journalists," in *2017 IEEE European Symposium on Security and Privacy (EuroS&P)*. Los Alamitos, CA, USA: IEEE Computer Society, apr 2017, pp. 385–400.

[26] J. R. P. Mauriés, K. Krol, S. Parkin, R. Abu-Salma, and M. A. Sasse, "Dead on Arrival: Recovering from Fatal Flaws in Email Encryption Tools," in *The LASER Workshop: Learning from Authoritative Security Experiment Results (LASER 2017)*. USENIX Association, Oct. 2017, pp. 49–57.

[27] S. E. McGregor, E. A. Watkins, M. N. Al-Ameen, K. Caine, and F. Roesner, "When the Weakest Link is Strong: Secure Collaboration in the Case of the Panama Papers," in *26th USENIX Security Symposium (USENIX Security 17)*. Vancouver, BC: USENIX Association, Aug. 2017, pp. 505–522.

[28] A. Peterson, "Edward Snowden sent Glenn Greenwald this video guide about encryption for journalists. Greenwald ignored it." https://www.washingtonpost.com/news/the-switch/wp/2014/05/14/edward-snowden-sent-glenn-greenwald-this-video-guide-about-encryption-for-journalists-greenwald-ignored-it/, 2014.

[29] V. Roth, T. Straub, and K. Richter, "Security and usability engineering with particular attention to electronic mail," *International Journal of Human-Computer Studies*, vol. 63, no. 1, pp. 51–73, 2005.

[30] S. Ruoti and K. Seamons, "Johnny's Journey Toward Usable Secure Email," *IEEE Security & Privacy*, vol. 17, no. 6, pp. 72–76, 2019.

[31] S. Ruoti, J. Andersen, S. Heidbrink, M. O'Neill, E. Vaziripour, J. Wu, D. Zappala, and K. Seamons, ""We're on the Same Page" A Usability Study of Secure Email Using Pairs of Novice Users," in *Proc. CHI Conference on Human Factors in Computing Systems (CHI'16)*, 2016.

[32] S. Ruoti, J. Andersen, T. Hendershot, D. Zappala, and K. E. Seamons, "Private Webmail 2.0: Simple and Easy-to-Use Secure Email," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology, UIST 2016, Tokyo, Japan, October 16-19, 2016*, J. Rekimoto, T. Igarashi, J. O. Wobbrock, and D. Avrahami, Eds. ACM, 2016, pp. 461–472.

[33] S. Ruoti, J. Andersen, T. Monson, D. Zappala, and K. Seamons, "A Comparative Usability Study of Key Management in Secure Email," in *Proceedings of the Fourteenth USENIX Conference on Usable Privacy and Security*, ser. SOUPS '18. USA: USENIX Association, 2018, p. 375–394.

[34] S. Ruoti, J. Andersen, D. Zappala, and K. E. Seamons, "Why Johnny Still, Still Can't Encrypt: Evaluating the Usability of a Modern PGP Client," *CoRR*, vol. abs/1510.08555, 2015. [Online]. Available: http://arxiv.org/abs/1510.08555

[35] S. Ruoti, N. Kim, B. Burgon, T. W. van der Horst, and K. E. Seamons, "Confused Johnny: when automatic encryption leads to confusion and mistakes," in *Symposium On Usable Privacy and Security, SOUPS '13, Newcastle, United Kingdom, July 24-26, 2013*, 2013, pp. 5:1–5:12.

[36] S. Sheng, L. Broderick, C. A. Koranda, and J. J. Hyland, "Why Johnny Still Can't Encrypt: Evaluating the Usability of Email Encryption Software," in *Proc. 2nd Symposium on Usable Privacy and Security (SOUPS'06)*. ACM, 2006.

[37] Signal, "Signal Support - Is it private? Can I trust it?" https://support.signal.org/hc/en-us/articles/360007320391-Is-it-private-Can-I-trust-it-, visited 02/01/2021.
[38] C. Stransky, D. Wermke, J. Schrader, N. Huaman, Y. Acar, A. L. Fehlhaber, M. Wei, B. Ur, and S. Fahl, "On the Limited Impact of Visualizing Encryption: Perceptions of E2E Messaging Security," in *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. USENIX Association, Aug. 2021, pp. 437–454. [Online]. Available: https://www.usenix.org/conference/soups2021/presentation/stransky
[39] C. Stransky, O. Wiese, V. Roth, Y. Acar, and S. Fahl, "Companion Website - EMail Paper," https://publications.teamusec.de/2022-oakland-email/, 2021.
[40] O. Tange, "GNU Parallel - The Command-Line Power Tool," *;login: The USENIX Magazine*, vol. 36, no. 1, pp. 42–47, Feb 2011. [Online]. Available: http://www.gnu.org/s/parallel
[41] H. Tankovska, "Statista - Number of e-mail users worldwide from 2017 to 2024," https://www.statista.com/statistics/255080/number-of-e-mail-users-worldwide/, 2021.
[42] A. Ulrich, R. Holz, P. Hauck, and G. Carle, "Investigating the OpenPGP Web of Trust," in *European Symposium on Research in Computer Security*. Springer, 2011, pp. 489–507.
[43] WhatsApp LLC, "end-to-end encryption - WhatsApp," https://blog.whatsapp.com/end-to-end-encryption/?lang=en, visited 02/01/2021.
[44] S. Whittaker and C. Sidner, "Email overload: exploring personal information management of email," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1996.
[45] A. Whitten and J. D. Tygar, "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0," in *Proc. 8th Usenix Security Symposium (SEC'99)*. USENIX Association, 1999.
[46] P. Zimmermann, "PGP Version 2.6.2 User's Guide," https://web.pa.msu.edu/reference/pgpdoc2.html, Oct. 1994.

APPENDIX

A. S/MIME and PGP Parsing Errors

During the parsing of the mails, we encountered parsing errors for 0.09% of all emails.

While the error rate for S/MIME is consistently low (cf. Fig. 9), the error rate for PGP is higher (cf. Fig. 8). Overall, we encountered parsing errors for up to 1.17% of the S/MIME and 22.09% of the PGP emails per year. While we cannot provide in-depth insights into the parsing errors due to ethics and data protection concerns, Figures 8 and 9 show that most of the parsing errors for both S/MIME and PGP were caused by the five most active users of the respective years. For example, we saw most parsing errors for S/MIME in 2005 and PGP in 2004 where our data set included 1,367 S/MIME emails and 47 users and 1,521 PGP emails and 129 users. In this year, the three most active S/MIME users contributed all parsing errors for S/MIME and the five most active PGP users contributed 306 parsing errors (88.26% of all PGP parsing errors). Even more extreme is the peak for PGP email related parsing errors in 2004 which were mostly caused by a single KMail [17] user.

While parsing errors in general are disappointing since they add noise to our results, the fact that the vast majority of errors was caused by a small group of or even single users alleviates their negative effects. Our evaluation focuses on a per-user and not a per-email view. A small number of users whose emails could not be parsed, for example, only has a low impact on the overall data concerning the use of multiple email clients.

[17] cf. https://apps.kde.org/kmail2/

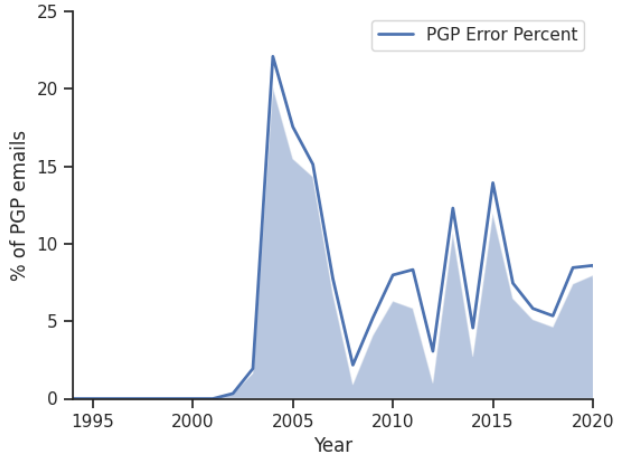

Fig. 8. PGP parsing errors over the years. The marked area is the parsing errors of the top 5 users of the year.

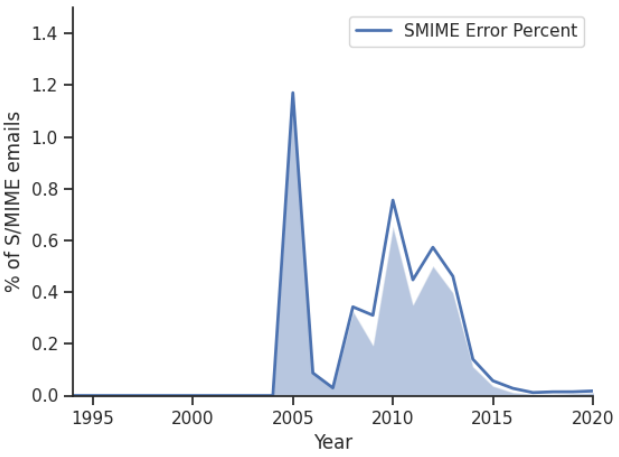

Fig. 9. S/MIME parsing errors over the years. The marked area is the parsing errors of the top 5 users of the year.

B. Pseudonymization Table

Table IV illustrates the pseudonymization we applied to the data.

C. Certificates issued for our university

Figure 10 shows the certificates that were issued by the DFN PKI in each year up until February 19, 2021.

Header-Data	Format
Message ID	SHA-256 with salt
User	SHA-256 with salt
User group	Categorized
Sender	SHA-256 with salt
Receivers	SHA-256 with salt
CC list	SHA-256 with salt
BCC list	SHA-256 with salt
Date	Bracketed into week
Client	Raw value
Folder	Categorized

S/MIME-Data	Format
Serial Number	SHA-256 with salt[1]
Not Valid Before and After	Bracketed into weeks[1]
Issuer Info	Raw value
Signature Algo	Raw value
Key size	Raw value
Key type	Raw value

PGP-Data[2]	Format
KeyID	SHA-256 with salt
Creation date	Bracketed into week
Expiration date	Bracketed into week
Type of key	Raw value
Length	Raw value
Key Algo	Raw value
Digest Algo	Raw value

[1] Applied to the leaf certificate
[2] Applied to both primary and sub key info

TABLE IV

COLLECTED METADATA FOR ALL EMAILS INCLUDING THE STORAGE FORMAT. FOR DATA PROTECTION REASONS AND IN CONSULTATION WITH OUR UNIVERSITY'S DATA PROTECTION OFFICER AND THE IT STAFF, WE CHOSE TO PSEUDONYMIZE SOME VALUES USING A HASH FUNCTION OR CATEGORIES.

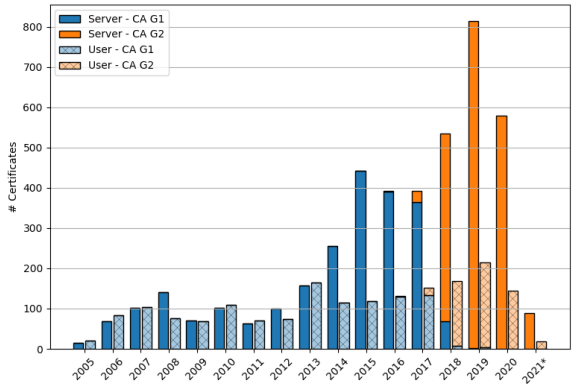

Fig. 10. X.509 certificates issued by the DFN PKI for our university from 2005 to 2020 using two different root certificates G1 and G2. G1 expired in 2019. Includes certificates issued until February 19, 2021.