Signatures for Privacy, Trust and Accountability in the Cloud: Applications and Requirements

Alaa Alaqra[1], Simone Fischer-Hübner[1], Thomas Groß[2], Thomas Lorünser[3], and Daniel Slamanig[4]

[1] Karlstad University, Sweden
[2] Newcastle University, UK
[3] Austrian Institute of Technology, Austria
[4] Graz University of Technology, Austria

Abstract. This paper summarises the results of a workshop at the IFIP Summer School 2015 introducing the EU Horizon 2020 project Prismacloud, that is, Privacy and Security Maintaining Services in the Cloud. The contributions of this summary are three-fold. Firstly, it provides an overview to the Prismacloud cryptographic tools and use-case scenarios that were presented as part of this workshop. Secondly, it distills the discussion results of parallel focus groups. Thirdly, it summarises a "Deep Dive on Crypto" session that offered technical information on the new tools. Overall, the workshop aimed a outlining application scenarios and eliciting end-user requirements for Prismacloud.

1 Introduction

Cloud computing is a very promising direction within ICT, but the practical adoption of cloud computing technologies may be greatly hindered by the lack of adequate technical controls to enforce the privacy of data and users in this outsourcing scenario. Solving this issues is especially challenging due to some fundamental properties of cloud computing such as being an open platform, its anytime and anywhere accessibility, as well as the intrinsic multi-tenancy, which introduce new security and privacy threats. In general, cloud computing is typically an outsourcing model and if the associated threats are not addressed adequately it leads to a tremendous risk for cloud users and their data. Thus, it is widely accepted that outsourcing of data and computations to third party cloud infrastructure requires various challenging security and privacy issues to be solved in order to gain users' trust. Besides the evident privacy and confidentiality issues associated with outsourced personal data and other type of confidential data (e.g., business secrets), which are a quite well understood albeit unsolved problem, this new computing paradigm introduces additional problems related to authenticity, verifiability and accountability. Basically, the question is how we can ensure that the cloud *works* as it is intended or claimed to do and how can the cloud be held accountable if deviations occur. Thereby, one may not only be concerned with the data itself, but also with processes (tasks/workflows) running in the cloud and processing the data. Moreover, such concerns may also be related to the used infrastructure itself.

Enforcing authenticity, verifiability and accountability for cloud based data processing by means of cryptography is one core topic within the recently started EU Horizon 2020 project PRISMACLOUD[5] on Privacy and Security Maintaining Services in the Cloud [11]. Its general goal is the research and development of tools and methodologies to strengthen the security, privacy and accountability for cloud based services, i.e., to make them more trustworthy. The main results will be showcased in different use-cases from the three application domains e-Health, e-Government and Smart Cities, which typically deal with sensitive data of citizens. To maximize the impact of the project results another focus of PRISMACLOUD is on the usability of developed solutions. Therefore, we are studying how users perceive such technologies if used within cloud based services, and elicit end user and human-computer interaction (HCI) requirements and guidelines for usable cryptography and protocols for the cloud. The aim is to design services which provide adequate security features but at the same time respect the users' needs in order to guarantee for the best acceptance of security technologies. The users should be able to understand and perceive the increased security and privacy they have when interacting with an augmented system while not being confronted with obstacles complicating their real tasks.

At the IFIP Summer School 2015 (Edinburgh, August 2015), the PRISMACLOUD project has organised a workshop comprising a series of parallel focus group sessions on the first day and a second-day "Deep Dive on Cryptography" workshop session. The motivation behind organizing this workshop related to the PRISMACLOUD project was as follows. Firstly, it was our intention to benefit from the knowledge of experts (from different domains) participating at the summer school in order to gather feedback, criticism and input on very early descriptions of the use-cases within PRISMACLOUD and to elicit end user and HCI requirements. Secondly, it was our aim to bring the attention of the audience to the cryptographic tools that are used and further developed within PRISMACLOUD. In particular, to attract interest from other researchers to also conduct research in this important field as well as interest from other security and privacy related research projects and researchers to cooperate with PRISMACLOUD. This paper summarises the content and discussion results of this workshop.

Outline. The remainder of this paper is structured as follows: In the next section, the cryptographic tools for the use-cases within PRISMACLOUD will be introduced. Section 3 briefly presents preliminary use-case scenarios in the areas of e-Health, e-Government and Smart Cities that are currently elaborated in PRISMACLOUD and helped explaining the ideas in the workshop and served as a basis for our focus group discussions. Section 4 presents the discussion results including the elicited end user and HCI requirements of five parallel focus groups that were part of the workshop. Section 5 will then summarise the results of the discussion on the second day on graph signatures and topology certification (as they have not been covered in the presented use-cases). Section 6 is finally rounding up this paper with overall conclusions.

[5] https://prismacloud.eu

2 Cryptographic Tools

Securing data over its life cycle in the cloud by means of cryptography is extremely challenging yet appealing to prevent many of provider related threats. This is due to the fact, that today widely used cryptography is designed to protect the confidentiality and authenticity of data in a very stringent way, i.e., without allowing for any modification. However, in the cloud setting it is important to support controlled altering and sharing of data in an agile way in order not to lose the benefits of cloud computing for cryptographically protected data. A very descriptive example is cloud storage. If we simply encrypt the data before uploading to the cloud we are protected from all major threats but completely lose the possibility to share or process the data, hence, we have to resign from almost all additional benefits of cloud computing for the sake of security. The same is true for authenticity protected data by means of signatures, every alteration of the data would immediately render the signature invalid, no matter how small it may be.

In Prismacloud we focus on the research and development of efficient cryptographic methods tailored to fit the needs of cloud computing and allow for controlled modification and sharing of data without giving up on the end-to-end security paradigm. We carefully selected technologies which have the potential to better protect the security of data during their stay in the cloud in a more agile way than currently possible. Subsequently, we briefly introduce some cryptographic tools that (a) are used within the use-cases presented and discussed in the workshop and (b) that have been presented throughout the second technical part of the workshop. Some of them do not appear in the use-cases that have been selected for the workshop and focus groups at the first workshop day. In particular, we will briefly present the concept of distributed cloud storage, different variants of signature schemes with special properties and the concept of (attribute-based) anonymous credentials.

Distributed cloud storage. Protecting the privacy, integrity and availability of stored data in the public cloud setting while at the same time allowing them to be shared in dynamic groups is a challenging problem. Currently, most cloud storage services store the data either unencrypted or apply encryption in a way that the keys remain under complete control of the cloud service provider. Hence, the data is susceptible to insider attacks and curious providers. In Prismacloud we follow a distributed systems approach and apply the cloud-of-clouds paradigm to increase availability and robustness. Here, the information is split into a number of shares [12], of which any subset of a fixed number allows the reconstruction of the original data. This approach is keyless and removes many obstacles in the area of usability and group key management [10]. Additionally it is capable to provide long-term security and everlasting privacy, which is very interesting for archiving of sensitive data. The confidentiality of data is guaranteed independently of the adversarial power and future developments, i.e., the rise of quantum computers. However, this assumption only holds as long as the majority of nodes in the cryptographic storage network have not been compromised. This assumption is different from conventional approaches and was a matter of discussion in the workshop.

Malleable and functional signatures. Malleable signatures are digital signatures that have some well-defined malleability property. This means, that signed data can be changed in a controlled way without invalidating the corresponding signature. In the following we will only very loosely discuss the two classes of malleable signatures that are of interest for the use-cases presented during the workshop. Firstly, malleable signatures that treat the signed message as structured data and allow to modify (e.g., black-out) well-defined parts of such a signed message. Such schemes, depending on their properties, i.e., who is allowed to perform the modifications, are modifications visible, etc., are denoted redactable [9,13] or sanitizable [1] signature schemes. The prime application of such a scheme is publishing a redacted version of a previously signed document where all sensitive information have been removed from the document without invalidating the original signature and thus the evidence for the authenticity of the document (cf. Figure 1).

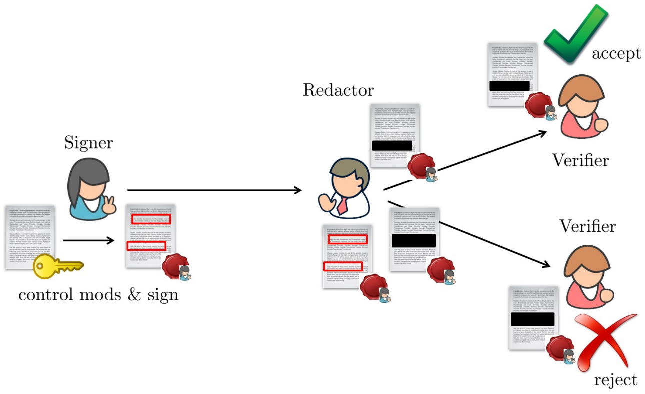

Fig. 1. Malleable signatures for document redaction.

Another class of schemes that usually treats messages as numeric data and targets on computing on signed data are called homomorphic signature schemes [6]. Basically, this means that there exists a public operation on signatures that carries over to the signed messages, e.g., one can compute the sum of single signed messages and derive a valid signature from the corresponding messages without requiring the secret signing key (cf. Figure 2). These schemes (and their practical efficiency) thereby greatly differ in the supported class of computations, e.g., linear functions, polynomial functions of some higher but fixed degree or arbitrary computations (fully homomorphic signature schemes).

Functional signatures [4] allow to delegate signature generation for message meeting certain conditions to other parties, who can then compute signatures

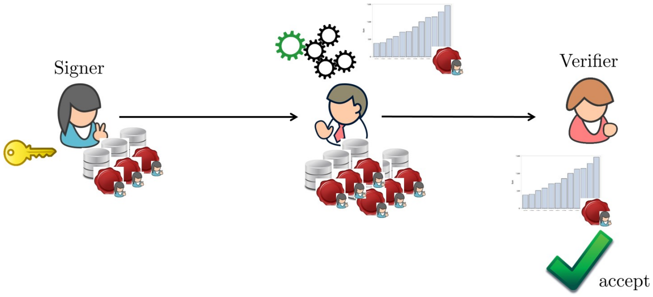

Fig. 2. Malleable signatures for numeric computations.

for a certain functionality on behalf of the original signer. Prime examples are proxy-signatures [3] for delegating signing capabilities and their application to certify computations on data (verifiable computations). Within PRISMACLOUD we want to study the application of aforementioned types of signature schemes to add verifiability features to data processing in the cloud in terms of end-to-end authenticity as well as verifiability of computations.

Graph signatures. Graph signatures [8] are a new primitive we investigate within PRISMACLOUD, which makes it possible that two parties engage in an interactive protocol to issue a signature on a graph. The resulting signature enables a prover to convince a verifier that the signed graph fulfils certain security properties (e.g., isolation or connectedness) without disclosing the blueprint of the graph itself. The foundational scheme for graph signatures [8] works on arbitrary undirected graphs. It encodes the graph data structure into a Camenisch-Lysyanskaya signature, making it accessible to zero-knowledge proofs of knowledge. The method for this is a form of Gödel numbering, that is, of representing data uniquely as products of prime numbers. This technique makes it possible that subsequent cryptographic proofs can argue over vertices, edges and labels. Within PRISMACLOUD we develop and optimize the use of graph signatures for practical use in virtualized infrastructures. Their application allows an auditor to analyse the configuration of a cloud, and to issue a signature on its topology (or a sequence of signatures on dynamically changing topologies). [7]. The signature encodes the topology as a graph in a special way, such that the cloud provider can prove high-level security properties such as isolation of tenants to verifiers. Furthermore, we will bridge between cloud security assurance and verification methodology and certification. We do this by establishing a framework that issues signatures and proves security properties based on standard graph models of cloud topologies and security goals stated in formal language, such that the virtualization assurance language VALID [2].

Anonymous credentials. Anonymous credentials (often denoted Privacy ABCs or simply ABCs) [5] are an important privacy-enhancing cryptographic tool that can be used to realize a privacy-friendly authentication mechanisms. In particular, it allows users to obtain credentials (that may contain various attributes of users) from some organization such that can later use them for authentication without the organization being able to track them. Moreover, if a user presents the credential more than once, these presentations cannot be linked together (unless special care is taken to allow such a mechanism). Finally, they allow data minimization. This means that the user does not need to reveal all the attributes encoded into a credential, but can selectively decide which attributes to show. Typically ABCs also allow a user to only prove that certain attributes satisfy some relation without revealing anything beyond, i.e., to demonstrate that the credential holder is older than some required threshold without revealing the birth date. In cloud based applications and services, the user's privacy is enormously endangered, since tracking user's data and behaviour is easily possible. Consequently, within Prismacloud we focus on bringing ABCs into practical application and also on improving their applicability.

3 Use-Case Scenarios

In the introductory workshop presentation, the following four use-case scenarios were presented to illustrate the use of the cryptographic tools of the project. The E-Health scenarios (a) and (b) use malleable signatures, the E-Government scenario is based on distributed cloud storage and the Smart Cities scenario involves anonymous credentials

(a) **E-Health: blood test.** Consider a case where a patient goes to the doctor for a routine check-up and takes an extensive blood test. The blood test is taken by the doctor's nurse and the results are uploaded to a cloud portal and are digitally signed by the nurse. The doctor has access to the complete blood test results. Later, the patient visits a dietitian, who requires few specific fields of the blood test. The patient doesn't want to reveal all fields from the extensive blood test. So the patient selects the mandatory fields from the extensive blood test for the dietitian to see and redacts ("blacks-out") the other fields.

Alternative case: Consider a case where the patient goes to the doctor for a routine check-up and takes an extensive blood test. The blood test results and diagnosis report are uploaded to a cloud portal and are digitally signed by the doctor. The doctor has access to the complete blood test results. However, the patient wants a second opinion from another doctor regarding her results. The patient doesn't want to reveal the diagnosis fields from the report. So the patient selects the blood test results for the second doctor while redacting ("blacking-out") the diagnosis field.

(b) **E-Health: smart phone monitor application.** Consider a case, where a patient has a smart phone training application that uses the sensors on the phone/wearable device to monitor and collect personal data of the patient. The patient would like to share only a statistical summary of activity progress information of the data collected by the application with her personal trainer without revealing sensitive medical data values.

(c) **E-Government: disaster files recovery.** For disaster recovery and backup purposes, IT providers of governmental institutions split their databases into multiple parts (shares) that are stored at independent cloud providers. Consider a case where a disaster occurs, and there is a risk of a potential data loss. To reconstruct data, only a predefined subset of shares stored at different cloud providers would be required, e.g., 4 shares out of 7.

(d) **Smart Cities: handicap parking.** Consider a case where handicapped citizens are required to use either their regular phones or smart phones to validate themselves in order for them to park at the handicap parking spot. Parking reservations are then stored centrally in the cloud for constantly monitoring the load of parking reservations. When using a regular phone, a control station by the parking will be used to authorize the parking using an SMS. When using a smart phone, the parking app would use the NFC badge (digital identification) and GPS location for authorization. With a privacy-enhanced solution based on a mobile phone based on an anonymous credentials-equipped mobile phone, the users could secretly authorise themselves for being eligible for this service without leaking any other information.

4 Day 1- Focus Groups Discussions

In the following subsections, we present a description of the workshop process. We give first an overview of the focus groups and then further details of the discussions and results per group in terms of the elicited requirements.

4.1 Workshop Format

A workshop in the form of expert focus group discussions was conducted on the first day with summer school participants who can be considered as experts in the field of privacy and security. The use-case scenarios in the areas of e-Health, e-Government, and Smart Cities developed in Prismacloud and briefly presented above in Section 3, were used in the workshop in order to give a context for the use of the project's cryptographic tools. The aim of the focus group discussions was to discuss use-case scenarios, to explore end user and HCI challenges of the scenarios and further elicit requirements in regards to usability, trust, and privacy.

The workshop consisted of informative and interactive parts. In total 25 participants with different research levels and backgrounds formed 5 interdisciplinary focus groups, coming mainly from Europe and Asia. The informative part consisted of a brief introduction to Prismacloud, the three use-case scenarios, and a technical overview of signatures schemes covering malleable and functional signatures and other Prismacloud crypto tools in preparations of the focus group tasks and discussions. Each group had a moderator (the authors of this paper) who guided the group through tasks, brainstorming activities, discussions, and feedback throughout the interactive sessions.

The interactive session consisted of three parts: (a) An introduction to the workshops agenda, materials, group forming, and group members' introductions.

(b) Selection of use-case scenarios to be discussed by that focus group and discussion of related cryptographic tools, and further the implications and features of those functions in regards to usability, privacy, and trust. (c) Requirements elicitation of cryptographic tools from part (b) to enhance usability, privacy, and trust in the cloud. For the brainstorming discussions, participants wrote short notes on opportunities and concerns that they see in regard to the selected case scenarios on post-it notes that were stuck on poster (see for example Figures 4 and 5 in the Appendix).

Results from the focus group sessions were documented as summaries by the moderator of each group. The summaries below followed the basic structure of:

A. Group participants
B. Use-case scenario
C. Key points of the discussion
D. Elicited requirements

4.2 Focus groups

The participants varied in formation of the 5 focus groups. For instance, one group consisted of only security and privacy PhD students group (FG1), others included a mix of security and privacy researchers with of participants with backgrounds in cryptography (FG2, FG3, FG4), cognitive science (FG2) and legal practice (FG5). For the use-case selection, the e-Health use-case scenario was chosen by FG1, FG3, and FG5, Smart Cities use-case was chosen by FG2, and FG4 discussed all. It was noted that the 5 focus groups have focused on different aspects of the scenario (which was expected), and the resulting requirements have reflected on these diverse focuses. All groups discussed the scenarios cryptographic aspects, however the focus was on control, privacy, and trust (FG1), functions, applications, and usability (FG2), rules and policies (FG3), cryptographic tools (FG4), as well as data types and legal rights (FG5), which is to be seen in the following subsections.

Focus Group No. 1:

(A) Group participants: The group consisted of 4 Computer Science PhD students doing research in IT Security & Privacy. A Computer Security professor and PRISMACLOUD project member acted as the workshop leader.

(B) Use-case scenario: As a scenario, the presented e-Health scenario on the redaction of blood test parameters in medical files stored in the cloud via malleable signatures was chosen and not further modified.

(C) Key points of the discussion: It was discussed that malleable signatures can in this case enhance privacy, as they give the data subject/redactor more control over what information to disclose to the verifier and what data she would like to redact. Hence, it allows the data subjects to enforce data minimization. At the same time, the barrier for patients to exercise control may be lower if they can do it electronically and thus directly, instead of having to request signed

redacted data offline (e.g. via mail). Patients may also put more trust into the health care provider, if they get options to control their data. Also, trust by the verifier can be enhanced, as the malleable signature guarantees that also the redacted document remains authentic.

However, increased patient control may also put extra burden and responsibility on the users. Moreover, it can also be debated whether patients should really have full electronic access to their medical dossiers, as they may not always be able to interpret all details and consequences correctly. From the patients perspective, they may not feel competent enough to do redactions themselves. For example, if they redact too much information, it may endanger their safety. They may therefore want to delegate this task to a trusted third party. However, accountability for the redaction may in this case be at stake.

As for redaction, doctors or nurses must be trusted to make competent decisions in regard to the amount of information that can be redacted by different patients considering both the patient's privacy and safety. If the redactor cannot be authenticated (i.e., in technical terms: the redaction operation is "unkeyed"), the verifier may lack trust in the redaction, e.g. may not be sure that really only information that was not needed in a certain context was redacted by authorized persons. Moreover, the patient may repudiate. If it is possible that the doctor can do the redaction and later claim that the patient did so, this may create privacy and trust issues.

If the signer who is in charge of sampling the blood test creates a malleable signature on the blood test that authorizing the patient concerned to do redactions on his blood test, then the identity of the patient may leak to the signer. However, for privacy reasons it is the practice that blood tests should be submitted anonymously.

It may affect trust if the verifiers cannot distinguish the cases when data has been redacted from documents or not. Also, privacy may be affected if the fact that information has been redacted (i.e. that the patient chose to hide certain medical values) cannot be hidden.

(D) Elicited requirements: The following list includes a number of requirements for enhancing privacy, trust and usability that were jointly suggested by the workshop participants:

- R1A It must be possible for the patient to delegate redactions to a specialist that he trusts; In this case, the delegate must be accountable for his actions.
- R1B The redactor should be accountable (i.e., the redaction operation should be a "keyed" operation).
- R1C Even if the redactor can be made accountable, there should be a possibility that the redactor can be anonymous or pseudonymous to the signer (so that the anonymity of blood tests can be guaranteed).
- R1D In dependence of the case, the redaction should be "visible" or "invisible" to the verifiers, i.e. in some cases the very fact that data was redacted should be hidden.
- R1E Usable guidelines and support are needed for informing users about how much information is advisable to redact taking both privacy and patient safety criteria into consideration.

– R1F The user interface should be based on suitable metaphors and HCI concepts and complementing tutorials for illustrating how the system works for promoting user trust in the claimed functionality of malleable signatures.
– R1G The definition of fields that can be redacted should follow the data minimisation principle while considering the patient's safety. Doctor and nurses need guidelines and support on how to define redactable fields while following these principles.

Focus Group No. 2:

(A) Group participants: The 5 participants of this group were 3 from computer privacy and security and 2 from cognitive science background. One issue regarding the mixture of the participants was related to their different levels of experience, which have hindered some discussion flows and interactivity, i.e. the two more senior researchers and practitioners in computer privacy and security were more dominant in the discussion due to their knowledge and expertise. An HCI Computer Security PhD. candidate and PRISMACLOUD project member acted as the workshop leader.

(B) Use-case scenario: When choosing the use-case, participants questioned the reason behind choosing a specific scenario and applicability of any chosen scenario. There was a discussion on how plausible the scenario is, and whether the scope is too narrow. Eventually, smart city and handicap parking was chosen as a preliminary case scenario.

(C) Key points of the discussion: The group started the discussion with the scenarios' functions. A main concern was raised on whether there is a need to use the cloud at all for this use-case scenario and when verifying credentials in the cloud which hardware and software to be considered from the users' side, in this case the discussion focused on the smart mobile phone. A debate arouse regarding whether the cryptographic tools are useful, it concluded with a suggestion to use attribute-based signatures to sign GPS coordinates as a claim of a handicapped person on a specific parking spot. Inspection measures versus linkability problem was brought up as there was a discussion on what is required to be considered and done in regards to this tradeoff, i.e., there is a need for inspection means, however linkability can't be avoided.

Some concern came up whether the application might give a false sense of privacy, where users might not be aware of the extent of data they are exposing. On the other hand, sabotaging users launching denial of service (DoS) and distributed denial of service attacks by anonymously reserving all parking places were discussed. Fraud and fault issues were addressed, and the discussion on how users can still lend out the handicap privileges despite the applications' main functions.

Finally, participants discussed usability issues with the app in comparison to the handicap card. The latter requires no effort on the behalf of the user, whereas the first is more demanding, i.e., credentials: there is a need for certain devices, and a level of understanding by the handicap users to get the application to work and show that the parking is authorized.

(D) Elicited requirements:

- R2A Trust requirements for the users: need of evaluators and transparency.
- R2B Each user must possess a credential that is securely stored on a mobile device, and a provably correct anonymous credentials protocol and implementation (validation + verification).
- R2C Important to protect the verifiers' availability and integrity (no corruption or coercion).
- R2D Payment requirement, even a little in order to mitigate DoS.
- R2E Revocation should be possible; temporary impaired/handicapped people (doctors/physicians can issue revocation).
- R2F Fraud inspection means are needed.
- R2G Usability: Less credentials to handle for easy decision making and less interferences with driving.
- R2H Suitable user interfaces and tutorials so that users can be aware of the systems functions and limitations.
- R2I Mobile application needs to be generic, for usability and appeal.

Focus Group No. 3:

(A) Group participants: The 5 group members consisted of a senior researcher in applied cryptography, a researcher and 2 PhD students in privacy and security related work within computing science, and a research engineer on privacy policies specification and their enforcement within a cloud computing environment.A technical Computer Security senior researcher and Prismacloud project member acted as the workshop leader.

(B) Use-case scenario: In this group, the e-Health scenario was chosen, and the discussion focused on the application of malleable signature schemes.

(C) Key points of the discussion: In particular, the discussion was about "blacking out fields" from medical data. In the beginning, there were some issues that needed clarification by the moderator (as there were questions from the participants which were only answered in the second part of the workshop). Afterwords, the discussion identified positive aspects of applying such schemes, e.g., more efficient processes (less interaction steps are required) and no longer requiring the signer if we want to give away authentic data to another party (offline feature). Nevertheless, the focus was more on the related problems and thus focused on what one would need to do in order to make such schemes applicable in practice.

It was identified that it is very important to specify redaction rules of how signed messages/documents are allowed to be redacted/modified. Thereby, it could be problematic if redacted versions of a document would be used in various different areas (e.g., e-Health and outside e-Health) - as this makes it hard to specify in which context which redaction is allowed. This could then lead to a redacted document that could be misused in the respective other area. Technically, one could counter this problem by using redaction policies (i.e., using a formal specification language to exactly specify what is allowed) and it should clearly (formally) define what is allowed to do in which context (it seems, however, that this is a highly non-trivial task). Policies could also support users (signers as well as redactors) to eliminate human errors and make such redaction tools easier to handle. Another problem that was identified in context of users is that users (signers) may not be able to comprehend what data to "mark" as being redactable. Consequently, it seems that for practical applications there is an inevitable need for policy and software support tool.

(D) Elicited requirements:

- R3A Important to specify redaction rules of how signed messages/documents are allowed to be redacted/modified.
- R3B Need for redaction policies (i.e., using a formal specification language to specify what is allowed) and it should clearly define what is allowed in which context.
- R3C Practical applications' strong need for policies and software support tools.

Focus Group No. 4:

(A) Group participants: The group consisted of (1) an associate professor of privacy enhancing protocols and privacy by design, (2) a principal research scientist in the Security and Cloud Research Lab with a focus on privacy enhancing technologies, accountability and the cloud, (3) a research engineer involved in developing a monitoring framework for cloud assurance and accountability, and (4) a PhD student working on data pseudonymization and anonymization. A technical Computer Security senior researcher and PRISMACLOUD project member acted as the workshop leader.

(B) Use-case scenario: In general, the group attempted to analyze all scenarios, but discussion got caught up on signatures. It started with detailed explanation of redactable signatures and the health use-case. The use of malleable signatures and verifiable computation in the blood test use-case was then discussed.

(C) Key points of the discussion: The discussion was focused on the tools. First the redactable signatures were introduced and explained by the moderator, the group understood the features and also the need for redaction in some situations, e.g. anonymous data sharing in health care applications, although they doubted the feature of anonymity, because inferences can be made by learning metadata. There are maybe better or additional means necessary like anonymization/pseudonymization to provide protection against re-identification. It was also a questions to which extent this features could be limited to third parties and selectively delegated. Another concern regarding the redactability was, if it was really deleted. This comment was also referring to the problem of re-identification. One participant questioned the use of malleable signatures, and claimed that the

concept is very close to ABC which even provide unlinkability and most of the features of redactable signatures cloud be implemented by the use of ABCs. He was interested in the advantage of redactable signatures compared to ABCs.

In the discussion of malleable signatures and verifiable computation, confidentiality was pointed out to be a more critical issue than authenticity. There was doubt about the use-case and participants thought the introduction of a trusted third party is dangerous. A concern regarding the danger of the third party adding not the right values to influence the result to their own favor,and that this scenario only makes sense if the final signature can also be used to verify that the right values have been included in the computation.

Reasons why ABC is necessary and what can be done with it were discussed. There was a concern regarding the smart city use-case with the electronic version of the disable batch. The fear was that it is still possible to link GPS or other metadata to anonymous credentials, e.g., license plate.

In the case of distributed storage, participants saw an opportunity to further compute data in such a setting which would be another advantage of such a system. However, they would like to see good technical arguments to make sure that they don't collaborate, because otherwise they would not fully trust this assumption to be true in many situations.

(D) Elicited requirements

- R4A Different scenarios for redactor roles are needed; if redactor=user, then use ABCs.
- R4B Need for proactive measures for introducing redactable fields.
- R4C Address the need for third parties, and improve means for trusting them (confidentiality).
- R4D Need for additional means to protect against re-identification and aid anonymization and pseudo anonymization.
- R4E Need for good technical arguments for trusting distributed storage systems.
- R4F Need to address scalability, what if many fields should be redactable.

Focus Group No. 5:

(A) Group participants: Five participants from computer security (2), privacy (1) and legal (2) background. The different backgrounds made for an interesting and inspiring discussion with multiple angles covered. The discussion was fruitful, albeit straying from the initial agenda. A technical Computer Security senior researcher and Prismacloud project member acted as the workshop leader.

(B) Use-case scenario: The discussion gravitated around finding scenarios for case studies, yet touched upon general principles. The scenario became the catalyst of the discussion, which yielded further considerations in multiple topics. As scenario, the group proposed e-Health as general area and specifically a fitness app that stores the data in the cloud. The question was raised what data is shared or stored locally on the user's device.

(C) Key points of the discussion: A core topic discussed by the group is the data types that need consideration in such a scenario, where the group named the following types:

- Medical data,
- Personal identifiable data,
- Location data,
- Time data (history over time), and
- Metadata (data about data).

The group raised the question about derivative data, i.e., data derived from the user's primary data, e.g., information learned and stored in Machine Learning Models. The question of ownership arises for the ownership and the user's rights with respect to that data. How could cryptography offer a chain of custody for such derivative data?

The group considered the overall risk of the scenario. Here, the opinion was voiced that having data stored in the cloud is equivalent to a risk. Further, it was raised whether the data should be stored in the cloud at all, and whether the benefits thereof make up for the risk. Further the group questioned the aggregation of data over time. What can parties learn from the user's cloud-stored data over the user's lifetime? Poignantly put, the question was asked "Will I get problems in 10 years time?" With respect to the data types mentioned before, the question was asked whether generally, there is too much data shared.

The primary question asked how cryptography can increase the trust in the system. To gain efficient solutions, the group advocated a "trust-but-verify" approach, which entails that one trusts parties optimistically, yet verifies that they are well behaved. This trust required that data processing goes beyond informed consent.

The group voiced the opinion that the legal system has an important role to play to ensure the privacy of the overall solution. Poignantly, this was put as "Court counters Curiosity". Furthermore, the question was raised whether it should be a human right to have access to cryptography. This discussion is in the context of privacy being supported by human rights and constitutions. Cryptography is a means to ensure privacy protection.

(D) Elicited requirements: The group discussed requirements on the system vis--vis of requirements on cryptographic primitives.

- R5A Simplicity and enhanced user experience.
- R5B Restrictions on retention of the primary data as well as on the retention of derived data.
- R5C Use of sticky policies (privacy policies attached, sticking, to the data) that enable a cross-system tracking of privacy policies, obligations and purpose-binding made. The sticky policies should be enforced by cryptography.
- R5D Strong purpose-binding throughout, that is, it is specified and enforced what purpose data can be used for. Purpose-binding could be enforced by encryption, e.g., attribute-based encryption with purpose credentials as attributes.

– R5E Need for misuse detection.

As cryptography requirements, the group advocated

– R5F Encryption of data at rest as minimal requirement, with the key stored on a user's device.
– R5G Malleable signatures should be used to allow the discovery of misbehaviour.
– R5H The cryptographic primitives employed should yield evidence.

5 Day 2- Deep Dive on Cryptography: Graph Signatures and Topology Certification

This section summarises the content and discussion for the second day "Deep Dive on Cryptography" workshop on graph signatures and topology certification, as this concept was not covered by the use-cases presented in the workshop. The objective of this second day workshop was in contrast to the first day focus groups not primarily the elicitation of requirements, but rather in addition to giving a short tutorial, the discussion of further possible application scenarios. The workshop contribution set the state illustrating that graphs are indeed a common data structure in computer science, naming examples of

– Social network graphs for a Blackhat organisation,
– Causality graphs (structured occurrence nets) for criminal investigations, and
– Topology graphs of virtualised infrastructures.

It was observed that often in these cases the integrity data substrate and the derived graph is not guaranteed and that the graphs contain confidential or sensitive information. Hence, there is a tension between integrity and confidentiality requirements.

Concretely, these conflicting requirements were illustrated for multi-tenant virtualised infrastructures, in which tenants seek to gain security assurance on the infrastructure while infrastructure providers (and other tenants) want to keep the blueprint of their infrastructure confidential. This problem is aggravated as a tenant's sub-system can be impacted by configuration changes elsewhere in the infrastructure. For instance, a misconfigured VLAN identifier elsewhere could lead to other tenants getting access to a tenant's private network, causing an isolation breach.

Naturally, tenants have little reason to trust the provider's assertions of the secure configuration of the entire infrastructure. They would require evidence for the security assurance based on an independent trust root. Hence, we introduce an auditor as third party, who inspects the low-level configuration of the infrastructure, derives a graph representation, and signs this representation. The signature is done in such a way that the provider can subsequently prove to the tenants that security properties they require are fulfilled. For instance, a tenant A could require that no other tenant has access to A's resources.

A more elaborate version of this scenario was presented in [7]. Figure 3 depicts the system model for the topology certification. The auditor continuously inspects the low-level configuration and issues multiple signatures for defined time instances. The provider receives all these signatures together with diff-logs on the graph representation. Henceforth, the provider is enables to prove that security properties on the topology are fulfilled for times asked about by the tenants.

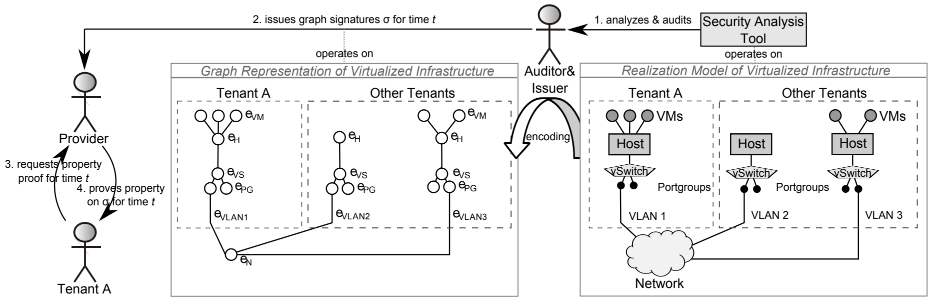

Fig. 3. System model of the topology certification proposal (from [7]).

It was pointed out that the graph signature primitive is generic, as it makes it possible to establish a signature on a graph independent of the question subsequently asked about the graph. The scheme is also expressive, as it can encode statements from arbitrary **NP**-languages.

The audience made observations that the graph signature scheme could be used for a variety of scenarios. One proposal made was that graph signatures could be used to prove that the surveillance and investigation of a secret agency has only infiltrated people with k degrees of distance in the social network graph of a designated target, based on a selector. Legal oversight requires such organisations to limit their investigations to a low number of hops from the designated selector. Previously, it was impossible to verify claims that the secret service agency has been compliant with the regulations. However, an independent auditor could derive a social network graph representation on surveillance requests and issue a graph signature, which would in turn enable the secret service agency to prove in zero-knowledge that it was compliant.

6 Conclusions

The workshop took advantage of diverse discussions that happened in the focus groups and workshop sessions for eliciting requirements for Prismacloud. Experts, coming from different areas and working backgrounds, have discussed opportunities and challenges in regards to enhancing privacy and trust in the cloud throughout the selected use-case scenarios discussions. It was concluded that the main notions to ensure trust are accountability, transparency, verification and authentication. There is a clear need for means, such as crypto tools, for enhancing users' privacy and control especially when dealing in different data types, such as (explicitly and implicitly) disclosed and derived data, in the cloud. Specific considerations are needed for Prismacloud, such as for redaction rules and policies which need to be clearly stated, e.g., in regard to the competence of the redactor and to the awareness of responsibilities associated with redaction, and delegation policies; i.e., delegations of redactions to a third party/specialist needs to fulfil trust requirements set by guidelines, policies, and laws. Privacy enhancing means by cryptography need usable guidelines and suitable interfaces and metaphors to communicate privacy incentives and risks to the users and ensure that a certain level of awareness is reached when using these means. Support from the legal perspective is necessary, e.g. by enforcing the encryption of users' data in the cloud as a requirement for privacy or, as one focus group discussed, even by establishing a human right to access cryptography.

Acknowledgements: The authors have received funding from the the European Union's Horizon 2020 research and innovation programme under grant agreement No 644962 (Prismacloud project).

References

1. Ateniese, G., Chou, D.H., de Medeiros, B., Tsudik, G.: Sanitizable Signatures. In: ESORICS. pp. 159–177 (2005)
2. Bleikertz, S., Groß, T.: A Virtualization Assurance Language for Isolation and Deployment. In: POLICY. IEEE (Jun 2011)
3. Boldyreva, A., Palacio, A., Warinschi, B.: Secure Proxy Signature Schemes for Delegation of Signing Rights. J. Cryptology 25(1), 57–115 (2012)
4. Boyle, E., Goldwasser, S., Ivan, I.: Functional Signatures and Pseudorandom Functions. In: PKC. pp. 501–519 (2014)
5. Camenisch, J.: Concepts around privacy-preserving attribute-based credentials - making authentication with anonymous credentials practical. In: Privacy and Identity Management for Emerging Services and Technologies. pp. 53–63 (2013)
6. Catalano, D.: Homomorphic Signatures and Message Authentication Codes. In: SCN. LNCS, vol. 8642, pp. 514–519. Springer (2014)
7. Groß, T.: Efficient certification and zero-knowledge proofs of knowledge on infrastructure topology graphs. In: Proceedings of the 6th edition of the ACM Workshop on Cloud Computing Security (CCSW 2014). pp. 69–80. ACM (2014)
8. Groß, T.: Signatures and efficient proofs on committed graphs and NP-statements. In: 19th International Conference on Financial Cryptography and Data Security (FC 2015). pp. 293–314 (2015)
9. Johnson, R., Molnar, D., Song, D., Wagner, D.: Homomorphic Signature Schemes. In: CT-RSA. pp. 244–262. LNCS, Springer (2002)
10. Lorüenser, T., Happe, A., Slamanig, D.: ARCHISTAR: Towards Secure and Robust Cloud Based Data Sharing. In: IEEE 7th CloudCom 2015, Vancouver, November 30 - December 3. IEEE (2015)
11. Lorünser, T., Rodriguez, C.B., Demirel, D., Fischer-Hübner, S., Groß, T., Länger, T., des Noes, M., Pöhls, H.C., Rozenberg, B., Slamanig, D.: Towards a New

Paradigm for Privacy and Security in Cloud Services. In: Cleary, F., Felici, M. (eds.) Cyber Security and Privacy, Communications in Computer and Information Science, vol. 530, pp. 14–25. Springer International Publishing (2015), http://dx.doi.org/10.1007/978-3-319-25360-2{_}2

12. Shamir, A.: How to share a secret. Commun. ACM 22(11), 612–613 (1979)
13. Steinfeld, R., Bull, L.: Content Extraction Signatures. In: ICISC. Springer (2002)

Appendix: Examples of Focus Group Notes

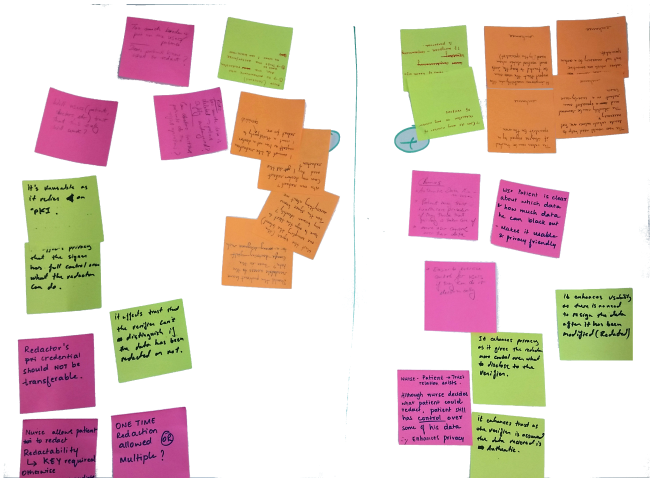
It's unusable as it relies on PKI.
privacy that the signer has full control over what the redactor can do.
it affects trust that the verifier can't distinguish if the data has been redacted or not.
Redactor's ~~pri~~ credential should NOT be transferable.
Nurse allows patient to redact
Redactability
↳ KEY required
Otherwise
ONE TIME Redaction allowed OR Multiple ?
User Patient is clear about which data & how much data he can black out
- Makes it usable & privacy friendly
it enhances usability as there is no need to resign the data after it has been modified (Redacted)
It enhances privacy as it gives the redactor more control over what to disclose to the verifier.
Nurse - Patient → Trust relation exists.
Although nurse decides what patient could redact, patient still has control over some of his data
∴ Enhances privacy
it enhances trust as the verifier is assured the data received is Authentic.

Fig. 4. Brainstorming notes on opportunities and concerns by focus group No. 1.

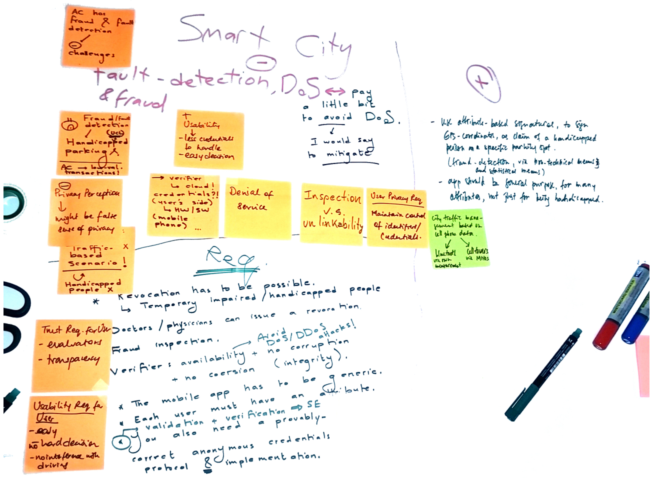
Smart City
AC has Fraud & fault detection
challenges
fault-detection, DoS & fraud
↔ pay a little bit to avoid DoS.
I would say to mitigate
Fraud/Fault detection
Handicapped parking
AC → low-val. transactions!
+ Usability
- less credentials to handle
- easy decision
Privacy Perception
might be false sense of privacy
→ verifier
↳ cloud!
credentials?! (user's side)
↳ HW/SW (mobile phone) ...
Denial of Service
Inspection v.s. unlinkability
User Privacy Req
Maintain control of identifiers/ Credentials.
City traffic management based on cell phone data.
bluetooth via own measurement
cell towers via MNOs
Traffic-based Scenario !
Handicapped People
- use attribute-based signatures, to sign GPS-coordinates, as claim of a handicapped person on a specific parking spot.
(fraud-detection, via non-technical means and statistical means)
- app should be general purpose, for many attributes, not just for being handicapped.
Req.
Revocation has to be possible.
* ↳ Temporary impaired / handicapped people
Doctors / physicians can issue a revocation.
Fraud inspection.
Verifier: availability + no corruption (integrity).
→ Avoid DoS / DDoS attacks!
+ no coercion
* The mobile app has to be generic.
* Each user must have an attribute.
validation + verification ⇒ SE
You also need a provably-correct anonymous credentials protocol & implementation.
Trust Req. for User
- evaluators
- transparency
Usability Req for User
- easy
no hard decision
- no interference with driving

Fig. 5. Brainstorming notes on opportunities, concerns an requirements by focus group No. 2.