AUTHENTICATION PROTOCOLS AND PRIVACY PROTECTION

AUTENTIZAČNÍ PROTOKOLY A OCHRANA SOUKROMÍ

ZKRÁCENÁ VERZE PH.D. THESIS

Obor: Teleinformatika
Oponenti:

Datum obhajoby:
KEYWORDS
Privacy, authentication, digital identity protection, cryptography, anonymity, smart-cards.

KLÍČOVÁ SLOVA
Soukromí, autentizace, ochrana digitální identity, kryptografie, anonymita, smart-karty.

Disertační práce je k dispozici na Vědeckém oddělení děkanátu FEKT VUT v Brně, Technická 10, Brno, 616 00

© Hajný Jan, 2012
ISBN 80-214-
ISSN 1213-4198
CONTENTS

Introduction 5

1 Current State Analysis 7
  1.1 Existing Techniques for Revocation 7
  1.2 Existing Attribute Authentication Systems 8

2 Thesis Objectives 11

3 Cryptography and Notation 12
  3.1 Used Cryptographic Primitives 12
     3.1.1 Commitment Schemes 12
     3.1.2 Σ-protocols 13
     3.1.3 Okamoto-Uchiyama Trapdoor One-Way Function 19
     3.1.4 Bao’s Verifiable Encryption 19
  3.2 Notation 20

4 Novel Scheme for Attribute Authentication 21
  4.1 Additional Contribution 22
  4.2 General Overview 23
  4.3 Cryptographic Design 24

5 Conclusion 29

Bibliography 31
User authentication is a service crucial for many electronic transactions. Without a secure verification of users, it would be impossible to provide many services both on the Internet and during everyday life. We need authentication services for the verification of identities and persons’ authorizations. The most common examples of systems, where authentication is a fundamental service, are the electronic banking systems, information systems or physical access systems. In addition to these fundamental systems, there are many non-critical applications which are used on daily basis like employee ID cards, library cards or discount cards. With the steep increase of the number of services provided electronically, it is reasonable to expect a strengthening demand for secure and reliable authentication systems. On the other hand, it is not only the security of service providers what is needed to protect. The security and privacy of users must be also protected. It is very important to keep in mind the fact that users release a lot of personal information by using authentication services. Every time we use an authentication system to get an access to a service, we release our identity, which can be abused by the service providers for tracking our behavior, profiling our usage of the service or even for impersonation. That is the reason why modern authentication systems provide privacy enhancing features protecting users’ identities and private data. In this thesis, the aspects of existing privacy enhancing authentication systems are analyzed and a new authentication scheme providing advanced features for privacy protection is developed.

The classical authentication systems, like RADIUS, Diameter or Kerberos, are frequently used for the identity verification. Based on the identity of a user, the system usually decides about the authorization. Therefore, the primal goal of these systems is to provide a user with some secret information and then do the verification of the possession of that secret. This approach is used in most present authentication systems because it is simple and allows many variations based on many forms of users’ secrets. The user passwords, secret keys, private asymmetric keys or even biometric data can be used. The classical authentication systems provide a relatively safe way to verify the identity of a user. On the other hand, underlying cryptographic primitives are usually simple and don’t allow building mathematical proofs of security. That is the reason why more complex authentication protocols have been developed in the end of 90’s.

The advanced authentication protocols allow a deeper mathematical analysis. These protocols are usually based on assumptions about underlying cryptographic primitives. In authentication systems, the use of so called Zero-Knowledge primitives have become very popular for the design of new protocols. These primitives allow building mathematical proofs of security. It is possible to prove that an authentication protocol based on Zero-Knowledge primitives releases exactly no more information about a user than it is designed for. This is very useful because the user can be then sure that no private information is released during the execution of the protocol. The Zero-Knowledge protocols are used in many practical systems today. Even advanced systems, which provide not only authentication but also more complex features, use the Zero-Knowledge cryptographic primitives. We describe the cryptographic background in Section 3 and build a new system from related primitives in Section 4.

Unfortunately, the verification of user identity brings some risks. Let us leave out the risk of stolen passwords or keys and let us consider the authentication protocols secure from this perspective. Even in that case, the service providers always learn the identities of users who are trying to use their
services. The identities can be considered privacy-sensitive information, from the privacy-protecting perspective it would be desirable that the service providers learn as little private information as possible. With the information about users’ identities, it is possible to track users, analyze their behavior and create user profiles. Based on this information, a more focused advertisement, prediction of user behavior or even the revelation of other personal information can be done. In this thesis, it is shown that more granularity to user authentication is needed. There are many services, like libraries, video archives or private Internet databases, where the identity disclosure is not necessary for authorization. In many cases, only the verification of some personal attributes (like age, license possession or citizenship) is sufficient for receiving a service. In these cases, disclosing identity is unnecessary and creates security risks. That is the reason why the attribute authentication systems have been introduced. In these systems, only a user-selected subset of private attributes can be disclosed. As a result, the users can stay anonymous during the use of services while their attributes are securely verified. These systems provide the maximum level of privacy protection - the users can reveal specifically only those private attributes which are needed by the service providers.

There are many reasons why the mentioned privacy preserving authentication systems are not yet used in commercial systems. Among others, it is the unwillingness of service providers who don’t want to provide their services to anonymous users. They want to protect their assets by being able to revoke invalid users and identify malicious users. Unfortunately, it is very difficult to achieve privacy and anonymity for honest users together with the ability to revoke and identify attackers. Currently, none of existing systems provides reliable and practical revocation of malicious users. Therefore, the goal of this thesis is to provide a cryptographic attribute authentication scheme with working revocation features.

In this thesis, the main focus is put on the privacy-preserving attribute authentication systems. They represent the most recent step in the development of authentication systems. To explicitly show the difference between classical authentication and attribute authentication, the following two definitions are included.

**Classical Authentication**: “Entity authentication is the process whereby one party is assured (through acquisition of corroborative evidence) of the identity of a second party involved in a protocol, and that the second has actually participated (i.e., is active at, or immediately prior to, the time the evidence is acquired).”

**Attribute Authentication**: Attribute authentication is the process whereby one party is assured (through acquisition of corroborative evidence) of the possession of particular attributes (like age, citizenship or driving license ownership) of a second party involved in a protocol, and that the second has actually participated (i.e., is active at, or immediately prior to, the time the evidence is acquired).

Chapter 1 describes the main competitors of the newly proposed cryptographic scheme which are the U-Prove of Microsoft and Idemix of IBM. Chapter 2 defines the objectives of the thesis. The requirements on the new authentication systems are stated here. The main output of this thesis, the new attribute authentication scheme, is described in Chapter 3. The information about the implementation on a smart-card platform, performance results and formal security proofs are provided in the full version of the thesis. The conclusion and proposals for future development are given in Chapter 5.
1 CURRENT STATE ANALYSIS

In the first part of this chapter, the existing revocation techniques are analyzed and their weaknesses are described. To author’s best knowledge, there is currently no scheme providing practical, multi-level revocation which is implementable on smart-cards.

In the second part of the chapter, the analysis of existing attribute authentication systems is provided. Currently, there are only two schemes which provide attribute authentication capabilities and have the potential to become widespread. They are the U-Prove and idemix, both analyzed in this chapter. These systems, both supported by large IT companies, are compared with the proposal. The weaknesses and missing features, which make them less practical, are identified.

1.1 Existing Techniques for Revocation

Revocation, a feature crucial for both attribute issuers and service providers, allows an authentication token to be revoked after its misuse, expiration, theft or loss. By omitting revocation, there is no mechanism for revoking malicious users, stolen verification tokens or lost verification smart-cards. Missing revocation is a strong reason why the service providers do not accept existing systems. It is too dangerous for them to provide their assets without being able to revoke invalid users, charge attackers for potential damage or identify clients who use stolen identities. The problem of missing revocation has been unsuccessfully addressed by many authors in papers dealing with privacy enhancing cryptography. A short overview of revocation techniques used in existing systems is provided there together with reasons why they were found impractical.

Blacklisting of Token IDs

Some systems, for example U-Prove [39], use a token identifier embedded to each transaction. The identifier is a public, unique and unchangeable number linked to the token. This number can be used to revoke the token by putting it on a blacklist. Nevertheless, this approach destroys unlinkability (the unique token identifier creates a link among all user’s verification sessions). Furthermore, a token can be revoked only by verifiers who already saw the token before and there is no mechanism for revoking tokens by their issuers. That is why the issuers have no power to revoke invalid, stolen or expired tokens.

Blacklisting of Secrets

The technique for blacklisting of secrets, used, for example, in [21], allows an invalid token to be revoked by using the knowledge of secret keys used for its construction. This technique can be used in cases where secret keys of users are revealed and for example made public on the Internet. In that case, a revocation authority can create a blacklist of these keys to prevent verifiers from accepting tokens based on leaked keys. Nevertheless, this technique works only if the user secrets are revealed. But in most cases, the secret keys never leave a protected device (like a smart-card), therefore they cannot be revoked. Moreover, lost, stolen or expired tokens (e.g., stored on a smart-card) cannot be revoked because their secret keys never become public.
Epochs of Lifetime

Epochs of lifetime are the official revocation technique of idemix [29]. Here, a credential carries an epoch of validity as a special attribute. In this case, the verifier can check whether the credential is fresh. The user is required to renew his credential for every new epoch. The disadvantage of this mechanism is that the revocation of credentials is never immediate, the revoker must wait until their expiration and the issuer must stop issuing new credentials to revoked users. The second major disadvantage is that the user must periodically run the issuance protocol with the issuer (or designated entity) to update his credential.

Accumulator Proofs

The most recent technique, analyzed in [37], allows both issuers and verifiers to revoke tokens immediately by publishing so called whitelists. In this technique, a user must provide a proof that his token is included on a list of valid tokens. This can be done anonymously and efficiently using so called accumulators which accumulate all non-revoked users. Most efficient techniques are based on bilinear maps [25, 40, 43, 25]. The disadvantage of these solutions is that the user must update his secrets every time any other user is revoked from the system. This is not a big problem when the user uses an on-line computer for his verification. On the other hand, if the user uses only an off-line device, like a smart-card, then he is unable to update his secrets. Therefore, the user is unable to use his token after some other users are revoked from the system. The objective of this thesis is to provide a system which can be used for everyday verification in libraries, pubs or hotels, therefore the smart-card implementation is crucial. That is the reason why the accumulator based techniques are considered impractical.

Verifiable Encryption of Secrets

The user identity or personal secrets can be encrypted inside a token in such a manner that only a trusted authority can do the revocation or identity disclosure by decryption. In this case, the system might be considered insecure from the perspective of a user who does not fully trust the authority. In fact, this is a likely problem since users would not welcome a scheme where a fixed third party can learn all information about their verification sessions, including their identities. In practical scenarios, the user would have no choice from more trusted authorities. This even more degrades his trust in such a dictated authority. Furthermore, there is a problem with unlinkability because the verifiable encryption must be randomized for each session, which might be inefficient. Even though mentioned in some papers, this technique has not been used for revocation in any well-known system for anonymous attribute authentication.

1.2 Existing Attribute Authentication Systems

There are two privacy preserving authentication systems which have the potential of getting largely widespread. They are the idemix of IBM [29] and the U-Prove of Microsoft [39].

Let us start with Microsoft’s U-Prove scheme developed by Stefan Brands [19]. Using U-Prove, a user can run the issuing protocol with an issuer to get a token with requested attributes. Then, the token can be presented to a service provider, who is able to verify the attributes and their ownership.
The user might decide to reveal only a subset of attributes. A hardware device might be enforced to participate during the presentation protocol. The scheme provides user anonymity, since the verifier is unable to identify a user who is presenting an attribute (if the user wishes so). Similarly, the issuer is unable to trace issued tokens. Nevertheless, the scheme does not provide the unlinkability of verification sessions, since one token cannot be spent more than once without spendings being linkable. Unfortunately, this makes the scheme vulnerable to attacks on users’ privacy. In many cases, just seeing what services are used together (at particular times) might be enough to de-anonymize the user. The recommendation of U-Prove authors is to use a different token for every verification session. This solution certainly works if the user has a computer connected to a network always available. In that case, he has enough resources to recompute tokens for every verification session. But this is not the case if we are using smart-cards only. In that case, the user cannot update tokens for each verification session because of the limited number of card’s memory write cycles. Besides write cycles, the system should be autonomous and working with smart-cards only. That allows usage in places where user’s computer is unavailable (libraries, pubs, country borders etc.). U-Prove also does not provide efficient revocation. It is possible to revoke tokens using their IDs, nevertheless this destroys unlinkability and is not available for Issuers. Unofficial mechanisms \[20\] need on-line token updates, thus require users’ devices connected to the Internet and are computationally inefficient. The solution is also impractical for computationally slow devices. U-Prove does not seem to aim to the verification based solely on smart-cards. Unfortunately, many systems (e.g. \[49\] [50]) have this limitation too, because they are unable to provide a user with a long-term token which can be spent many times without being linkable.

The attacks using linking of verification sessions are not applicable to credential systems. These systems were introduced by Chaum \[36\], improved by Camenisch and Lysyanskaya \[26, 27\] and since then they have been redefined or improved many times to become idemix of IBM \[29\]. Similarly to U-Prove, these systems allow a user to be verified as a member of an organization, more generally as a holder of an attribute. This can be done efficiently, anonymously, untraceably a provably securely. There is also no need for token re-registration, computation delegation or communication with third party entities. The system has already been implemented in both computer and smart-card environment \[35\] \[21\]. Unlike U-Prove, idemix can run solely on smart-cards (JavaCards) without losing unlinkability. That is the reason why idemix can be implemented as an eID, where the verification
sessions cannot be linked together. Nevertheless, also idemix lacks some important features. The basic revocation of users is still not fully resolved, since the initial proposal [26] is too computationally inefficient for a smart-card implementation, more recent solutions require credential updates [37, 28, 29], are not supported by current smart-card platforms [10, 43] or both [25]. Furthermore, the idemix does not provide mechanisms for the identification of malicious users. This feature has been proposed, but is not implementable on smart-cards due to high complexity [26], or is focused rather on accountability and the fulfillment of a predefined contract [24]. Full identification of internal malicious users is considered very important in this thesis, since it protects service providers’ assets. It is supposed that any system lacking the ability to identify and charge malicious users cannot be widely accepted by commercial service providers. Moreover, the verification time of idemix implementation [21] is around 10 s for the basic version. Although it is a dramatical improvement in comparison to previous schemes, it can be still considered quite a long time for verification in a real life.

An overview of features supported by the proposed scheme and its main competitors, U-Prove and idemix, is provided in Table 1.1. In Speed row, the letter “u” represents the number of undisclosed attributes. Speed is measured by the number of exponentiations (exp). The comparison is based on information from the official U-Prove [39] and idemix [29, 21] specification.

\[JCOP\text{ }v2.2/41\text{, restricted attribute security model, no revocation, no malicious user identification.}\]
2 THESIS OBJECTIVES

The general objective of this thesis is to provide more privacy and identity protection in user authentication. Although it may sound contradictory to provide both privacy and authentication, there are systems which partially allow that. With attribute authentication systems, it is possible that users only prove the possession of some personal attributes (for example age, citizenship or driving license ownership) without disclosing their identity. More types of systems can be called attribute authentication systems. Therefore, the analysis of existing systems is the first objective of this thesis.

The most promising ones, credential schemes and anonymous tokens, are introduced in Chapter [1]. Unfortunately, these complex cryptographic systems do not provide all required features. Therefore, the main objective of this thesis is to design a new cryptographic scheme which will support all following privacy preserving features.

- **Security**: the novel scheme will be built using strong cryptographic primitives.
- **Anonymity**: the scheme will protect user’s identity by providing anonymity during the verification session.
- **Untraceability**: no entity in the scheme will be able to trace a particular user.
- **Unlinkability**: more verification sessions of a single user will not be mutually linkable to a single profile.
- **Selective disclosure of attributes**: users will be able to disclose and prove the possession of any subset of their private attributes.
- **Non-transferability**: user’s will be strongly discouraged by used cryptography from lending their attributes to other entities.
- **Practical, complex revocation**: invalid credentials will be revocable and malicious users will be traceable or even identifiable while honest users will stay anonymous.

The above enlisted privacy enhancing features are partially supported by existing systems. Nevertheless, two of these features are very difficult to provide. It is the unlinkability of verification sessions where it is necessary to make all sessions completely different and mutually unlinkable. All authentication data must be always regenerated and randomized to provide unlinkability. This operation can be too complex for low-performance devices such as smart-cards. The leading scheme for anonymous tokens, the U-Prove [39], currently does not support the unlinkability. The other privacy enhancing feature, revocation, is even harder to obtain. For revocation, some authority or authorities must be able to reveal all sessions of a particular user and invalidate them. This virtually contradicts anonymity, untraceability and unlinkability. Currently, satisfactory revocation is not supported by existing practical solutions. The new scheme for attribute authentication designed in this thesis will provide all required features, including unlinkability and revocation.

Furthermore, it is necessary to provide guarantees for service providers that the new system cannot be abused by users who want to exploit anonymity. Such guarantees are not provided by existing systems which leads to their rejection by the commercial sector. Therefore, a complex revocation of malicious user anonymity must be provided and users’ responsibility for their acts must be assured.

By providing all mentioned privacy-enhancing features, particularly the unlinkability of sessions and practical revocation, it will be possible to construct new practical authentication systems which will be protecting user privacy and digital identity much more than current solutions do.
3 CRYPTOGRAPHY AND NOTATION

This chapter describes the cryptographic constructions and the notation used in this thesis. The cryptographic primitives are described here because they are used in subsequent chapters describing the related work and the proposed scheme. Most cryptographic primitives use the concept of provable security. In this concept, mathematical models giving proofs about primitives’ features can be built. Thus, it is possible to build proofs about the security of protocols. The primitives used in this thesis are shown here together with related proofs and definitions. The notation is also described here because it is used not only in chapters devoted to the novel scheme design but also during the analysis of most related systems.

The first construction, cryptographic commitments, is used in cases where a user needs to commit to a number without disclosing it. Two versions of commitments are shown here, the computationally hiding and perfectly binding commitments and the perfectly hiding and computationally binding commitments. Then, the Σ-protocols are introduced. They represent an efficient and practical variant of ZK protocols. The Σ-protocols for proofs of knowledge, representation and discrete logarithm equivalence are described in next sections because they are very frequently used in Chapter 4. Using above mentioned primitives, more complex constructions can be built. The Okamoto-Uchiyama Trapdoor One-Way Function and Bao’s Verifiable Encryption are described as they are used in the design of the proposed scheme.

3.1 Used Cryptographic Primitives

Most cryptographic primitives described in this section exist in more variants. In this thesis, the variants working with modular arithmetics and groups where a discrete logarithm is hard to compute are used.

3.1.1 Commitment Schemes

A cryptographic commitment scheme can be used in scenarios where a user (U) is required to bind to a number without disclosing it. Therefore, there are two properties which must be fulfilled. They are the hiding property and the binding property. A practical example of using the commitment schemes is the situation where a user generates a random number $w$ and computes a commitment $c = \text{commit}(r, w)$ using a randomness $r$. Then, the user can disclose $c$ to some verifier (V). Although $c$ is disclosed, the verifier is unable to learn $w$ from it (hiding property) and the user is unable to change his $w$ without changing $c$ (binding property). More formally:

- **Hiding property**: for V, it is difficult\(^1\) to learn $w$ from $c$.
- **Binding property**: for U, it is difficult to compute $(r', w')$ such that $c = \text{commit}(r, w) = \text{commit}(r', w')$ and $w \neq w'$.

Either of these properties can hold unconditionally while the other holds computationally. In commitment schemes, it is impossible to have both hiding and binding properties holding unconditionally. Two examples of commitment schemes are described here: the Discrete Logarithm (DL) commitment and the Pedersen commitment \(^{40}\).

\(^1\)It is either impossible or computationally infeasible.
Discrete Logarithm Commitment Scheme

The DL commitment scheme works with the subgroup $\mathbb{Z}_q^*$ of a multiplicative group $\mathbb{Z}_p^*$. The subgroup $\mathbb{Z}_q^*$ is defined by a generator $g$ of order $q$ in mod$p$, where $q$ and $p$ are large primes and $q$ divides $p - 1$. The same settings is used in the Digital Signature Algorithm (DSA) [41]. Numbers $g, q, p$ are system parameters which are made public. To commit to a number $w < q$, a user computes $c = g^w \mod p$. The user can later decide to open the commitment by making $w$ public. The scheme is computationally hiding and unconditionally binding.

- **Hiding property**: for V, it is difficult to learn $w$ from $c$ because it is computationally hard to compute discrete logarithms in DSA subgroup [41]. Since the DL hardness assumption is used, the hiding property holds only computationally.
- **Binding property**: for U, it is impossible to compute $w'$ such that $c = g^w \mod p = g^{w'} \mod p$ because the order of $g$ in mod$p$ is $q$ and both $w$ and $w'$ must be smaller than $q$. Because no assumption is used here, the property holds unconditionally.

Pedersen Commitment Scheme

The systems parameters $g, q, p$, used in the DL commitment scheme, are also used in the Pedersen scheme [46]. Additionally, one more generator $h$ is used. It is important, that $\log_g h \mod p$ is unknown to the user. The commitment to a secret number $w$ is computed as $c = g^w h^r \mod p$ where $r$ is a random number chosen by the user and smaller than $q$. The user can later decide to open the commitment by making $(w, r)$ public. The scheme is unconditionally hiding and computationally binding.

- **Hiding property**: for V, it is impossible to learn $w$ because $r$ randomizes $c$ to any number in $\mathbb{Z}_q^*$. In other words, with fixed $c$, for each $w < q$ there is a unique $r$, such that $c = g^w h^r \mod p$ holds. That is why the hiding property holds unconditionally.
- **Binding property**: for U, it is difficult to find $(w', r')$ such that $c = g^{w'} h^{r'} \mod p = g^w h^r \mod p$. Computing $(w', r')$ is computationally equivalent to solving the DL problem which is assumed hard. That is why the property holds only computationally.

3.1.2 $\Sigma$-protocols

The $\Sigma$-protocols [22, 32] (Sigma-protocols) provide similar properties as the full ZK protocols while staying practical. They can be used in more complex constructions as building blocks. In this thesis, the $\Sigma$-protocols are used in the cryptographic core of the system proposed in Chapter 4. The $\Sigma$-protocols have following properties.

- **3-way pattern**: there are 3 messages, one goes from the Prover to the Verifier, the second (the challenge) goes vice versa and the last one (the response) goes again from the Prover to the Verifier. Besides these messages, system parameters must be pre-shared before running the protocol.
- **Completeness**: an honest Prover who knows secret information is always accepted by the Verifier.
- **Special Soundness**: Prover’s secret presence is required for the acceptance. That is assured by the existence of a probabilistic algorithm $M$, called knowledge extractor, which is able to
extract the secret from P if it is given a rewindable black-box access to it. This property assures that it is impossible to correctly answer more V’s requests without knowing the secret.

- **Honest Verifier Zero Knowledge (HVZK):** similarly to full ZK, this property guarantees that no secret information is released during the protocol run. In contrast to full ZK, the HVZK provides such security guarantee only for the communication with honest Vs. This might seem impractical but fortunately this problem can be resolved by slight modifications of the protocol. The Σ-protocols can be made computationally ZK using [34] and fully ZK using [23]. The HVZK simulator must have the same characteristics as the full ZK simulator, thus run in polynomial time and generate an output indistinguishable from the real protocol run.

Similarly to ZK protocols, the Σ-protocols can be used for proving statements. Unlike full ZK, these protocols cannot be used for proving all statements. The most practical statement proofs are based on the DL problem or the RSA problem. By using Σ-protocols, the prover can efficiently prove that he knows a discrete logarithm of some public value, a representation of some public value and that the discrete logarithm of some different public values are the same with respect to different bases. These practical protocols are the core of many modern cryptographic constructions. They form the basis of both the scheme presented in this thesis and its most related systems shown in Chapter 1.

All Σ-protocols presented in this chapter work with the DSA subgroup \( \mathbb{Z}_{q}^* \), namely with generators \( g_i \) of order \( q \) in modulo \( p \), where \( q, p \) are large primes and \( q \) divides \( p - 1 \). Parameters \( g_i, q, p \) are pre-shared to both Prover and Verifier.

### Proofs of Knowledge of Discrete Logarithm

The Proof of Knowledge of Discrete Logarithm (PKDL) protocol can be used by the Prover to give a proof about the knowledge of a discrete logarithm of some public value \( c \) with respect to generator \( g \) and modulus \( p \). Using PKDL, the Prover is able to convince V that he knows \( w = \log_g c \mod p \) without actually disclosing it. The practical protocol, called Schnorr protocol [48], is depicted in Figure 3.1.

The protocol is Zero-Knowledge (ZK) against honest verifiers; it can be made ZK against any V using [23]. It can run non-interactively using [34]. With some restrictions, the protocol can be used in groups with hidden order by sending answers in \( \mathbb{Z} \) [33]. The protocol is considered to be a Σ-protocol because it fulfills all requirements.

- **3-way pattern:** the Schnorr protocol is a 3-leg protocol with 3 moves.
- **Completeness:** an honest Prover knows \( w \), therefore he can compute an answer \( z = (r - ew) \mod q \), which is always accepted in the final check:

  \[
  \bar{c} \equiv g^{z}c^{e} \equiv g^{r-ew}(g^{w})^{e} \equiv g^{r}g^{-e}g^{ew} \equiv g^{r} \equiv \bar{c} \pmod{p}.
  \]  

  (3.1)

- **Special Soundness:** the knowledge extractor uses the standard rewinding technique, thus it asks P for a response on a challenge \( e \), then renews the Prover to step 1 and asks for a response on a different challenge \( e' \). Then, the extractor has two answers \((z, z')\) on different challenges
Fig. 3.1: Schnorr’s proof of knowledge of discrete logarithm protocol.

In the above described PKDL protocol, the Verifier sends a random challenge $e$ to the Prover. But the Verifier does not have to be honest in all cases and might send malicious non-random $e$ with the purpose of learning Prover’s secrets. To mitigate these attacks and to lower communication complexity, the non-interactive version of the protocol can be used. In that version, the request is replaced by the $|q|$-bit-long output of a hash function $H$ of previous protocol values and public system parameters. Such approach is known as Fiat-Shamir heuristics \[34\]. Assuming the hash function is secure, its output is always random to Prover as he is unable to predict its value due to one-wayness. The protocol security then depends on the security of the hash function. The non-interactive version
of the PKDL protocol is depicted in Figure 3.2. All proofs of knowledge can be translated from the interactive version to the non-interactive version using the Fiat-Shamir heuristics so we do not have to consider only honest verifiers but all verifiers.

**Proofs of Knowledge of Representation**

In cryptographic constructions, it is very common to compute multi-exponentiations in multiplicative groups. The Pedersen commitment is the typical example of multi-exponentiation where two generators are used. Using $k$ generators and exponents, the multi-exponentiation can be denoted as

$$ c = \prod_{1 \leq i \leq k} g_i^{w_i} \mod p, \quad (3.3) $$

where $(w_1, w_2, \ldots, w_k)$ is called the representation of $c$ with respect to generators $(g_1, g_2, \ldots, g_k)$ in mod $p$. With $(g_1, g_2, \ldots, g_k), q, p, c$ being public, the Verifier might ask the Prover to give a proof of knowledge of $(w_1, w_2, \ldots, w_k)$. In our proposal, a proof of representation for $k = 2$ is used, therefore an example for $k = 2$ based on the Schnorr protocol is given in Figure 3.3.

The protocol is the parallel composition of the former PKDL protocol so it retains its original properties. Similarly as the PKDL protocol, it can be transformed to a full ZK protocol and can run non-interactively. All requirements on Σ-protocols are preserved.

- **3-way pattern**: the protocol is a 3-leg protocol with 3 moves.
- **Completeness**: an honest Prover knows $(w_1, w_2)$, therefore he can compute answers $z_1 = r_1 - ew_1$ and $z_2 = r_2 - ew_2$, which are always accepted in the final check:

$$ \bar{c} \equiv g_{r_1} g_{r_2} c^e \equiv g_1^{r_1 - ew_1} g_2^{r_2 - ew_2} (g_1^{w_1} g_2^{w_2})^e \equiv g_1^{r_1} g_2^{r_2} g_1^{-ew_1} g_2^{-ew_2} g_1^{ew_1} g_2^{ew_2} \equiv g_1^{r_1} g_2^{r_2} \equiv \bar{c} \mod p. \quad (3.4) $$

- **Special Soundness**: the knowledge extractor uses the same technique as in PKDL, thus it asks P for responses on a challenge $e$, then rewinds the Prover to step 1 and asks for responses
Fig. 3.3: Proof of knowledge of representation protocol.

on a different challenge $e'$. Then, the extractor has two pairs of responses $((z_1, z_2), (z'_1, z'_2))$ on different requests $(e, e')$ with a fixed first step. From $((z_1, z_2), (z'_1, z'_2))$ and $(e, e')$, the extractor is able to compute the representation $(w_1, w_2)$.

\begin{align*}
1 & \equiv g_1^{z_1 - z'_1} g_2^{z_2 - z'_2} e^{e - e'} \pmod{p} \\
c & \equiv g_1^{z'_1 - z_1} g_2^{z'_2 - z_2} \pmod{p} \\
w_1 & \equiv \frac{z'_1 - z_1}{e - e'} \pmod{q} \\
w_2 & \equiv \frac{z'_2 - z_2}{e - e'} \pmod{q}
\end{align*}

(3.5)

- **Honest Verifier Zero Knowledge (HVZK):** the protocol simulator works in these steps:
  1. It takes $c$ and $e$ as inputs.
  2. It picks $z_1, z_2 \in R \mathbb{Z}_q$.
  3. It computes $\tilde{c} = g_1^{z_1} g_2^{z_2} e^e \pmod{p}$.

The output $(c, e, \tilde{c}, z_1, z_2)$ is indistinguishable from the real run of the protocol. The simulator runs in expected polynomial time.

The proofs of properties for $k > 2$ follow the same argumentation as the one provided above.

**Proofs of Discrete Logarithm Equivalence**

Using $\Sigma$-protocols, it is easy to give a proof that two different DL commitments $c_1, c_2$ were constructed using the same exponent $w$, so that $w = \log_{g_1} c_1 = \log_{g_2} c_2$. For this type of proof, the Proof of Discrete Logarithm Equivalence (PDLE) protocols can be used. In the next example, the Prover needs to prove that $\log_{g_1} c_1 = \log_{g_2} c_2$ using the same DSA subgroup settings like in previous examples. The user not only gives a proof of DL equivalence but also proves that he knows $w$. Similarly to previous examples, the system parameters $g_1, g_2, p, q$ and values $c_1, c_2$ must be pre-shared to $P$ and $V$. $P$'s secret input is $w$, the DL of $c_1, c_2$. The protocol is depicted in Figure 3.4.
\begin{align*}
\text{Prover} & \quad \text{Verifier} \\
& \quad c_1 = g_1^w \mod p, c_2 = g_2^w \mod p, g_1, g_2, q, p \\
& \quad r \in \mathbb{R} \mathbb{Z}_q \\
& \quad \bar{c}_1 = g_1^r \mod p \\
& \quad \bar{c}_2 = g_2^r \mod p \\
& \quad \bar{c}_1, \bar{c}_2 \quad \quad \quad \quad \quad \quad \quad e \in \mathbb{R} \mathbb{Z}_q \\
& \quad z = (r - ew) \mod q \\
\end{align*}

Fig. 3.4: Proof of discrete logarithm equivalence.

Similarly to both former protocols, also PDLE can run non-interactively using [34] and obtain full ZK using [23]. The argumentation for all properties required for \( \Sigma \)-protocols follows the former examples.

- **3-way pattern**: the PDLE protocol is a 3-leg protocol with 3 moves.
- **Completeness**: the honest Prover knows \( w \), therefore he can always compute an answer \( z = (r - ew) \mod q \) which is always accepted in the final checks.

\begin{align*}
\bar{c}_1 & \equiv g_1^z c_1^e \equiv g_1^{r-ew} c_1^e = g_1^r g_1^{-ew} g_1^w \equiv g_1^r = \bar{c}_1 \quad (\mod p) \\
\bar{c}_2 & \equiv g_2^z c_2^e \equiv g_2^{r-ew} c_2^e = g_2^r g_2^{-ew} g_2^w \equiv g_2^r = \bar{c}_2 \quad (\mod p) \\
\end{align*}

(3.6)

- **Special Soundness**: the knowledge extractor uses the standard rewinding technique, thus it asks \( P \) for a response on a challenge \( e \), then rewinds the Prover to step 1 and asks for a response on a different challenge \( e' \). Then, the extractor has two answers \((z, z')\) on different challenges \((e, e')\) with the first step fixed. Therefore, the extractor has 2 equations for each of 2 commitments which passed the check. The secret \( w \) can be computed, e.g., from the equation for \( c_1 \), similarly to the original PKDL protocol.

\begin{align*}
1 & \equiv g_1^{z'-z} c_1^{-e'} \quad (\mod p) \\
1 & \equiv g_1^{z'-z} c_1 \quad (\mod p) \\
c_1 & \equiv g_1^{z'-z} \quad (\mod p) \\
w & \equiv (z'-z) / (e - e') \quad (\mod q)
\end{align*}

(3.7)

- **Honest Verifier Zero Knowledge (HVZK)**: the simulator works in these steps:
  1. It takes \( c_1, c_2 \) and \( e \) as inputs.
Tab. 3.1: Overview of used cryptographic constructions.

<table>
<thead>
<tr>
<th>Construction</th>
<th>Group</th>
<th>Where Used</th>
<th>Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL Commitments</td>
<td>DSA</td>
<td>R</td>
<td>$c = \text{commit}(w)$</td>
</tr>
<tr>
<td>Proofs of Knowledge</td>
<td>DSA, OU</td>
<td>R</td>
<td>$PK{w : c = g^w}$</td>
</tr>
<tr>
<td>Proofs of Representation</td>
<td>DSA, OU</td>
<td>V</td>
<td>$PK{w, w' : c = g^w g'^{w'}}$</td>
</tr>
<tr>
<td>Proofs of DL Equivalence</td>
<td>DSA, OU</td>
<td>R, V</td>
<td>$PK{w : c_1 = g^w \land c_2 = g^2}$</td>
</tr>
<tr>
<td>OU Trapdoor OWF</td>
<td>OU</td>
<td>R, Rev</td>
<td>$c = g^w \mod n$</td>
</tr>
<tr>
<td>Verifiable Encryption</td>
<td>OU</td>
<td>V, Rev</td>
<td>$VE{x : c = g^x}$</td>
</tr>
</tbody>
</table>

Glossary:
OU: Okamoto Uchiyama [45]
DSA: Digital Signature Algorithm [41]
OWF: One-Way Function
R, V, Rev: Registration, Verification and Revocation protocol of the proposed scheme

2. It picks $z \in_R \mathbb{Z}_q$.
3. It computes $\bar{c}_1 = g_1^z c_1^e \mod p$ and $\bar{c}_2 = g_2^z c_2^e \mod p$.

The output $(c_1, c_2, e, \bar{c}_1, \bar{c}_2, z)$ is indistinguishable from the real run of the protocol. The simulator runs in expected polynomial time.

The PDLE protocol can be used not only for proving the equivalence of two logarithms but also for general proofs of equivalence of any exponents in different representations. Similarly to former protocols, PDLE can be used for proofs of DL in different groups [33]

3.1.3 Okamoto-Uchiyama Trapdoor One-Way Function

Let $n = r^2 s$ and $r, s$ be large primes. Pick $g \in \mathbb{Z}_n$ such that $g \mod r^2$ is a primitive element of $\mathbb{Z}_{r^2}^*$. Then $c = g^x \mod n$ is a trapdoor one-way function with $r$ as a trapdoor. Value $x$ can be computed using the trapdoor as $x = \frac{((c^{r-1} \mod r^2) - 1)/r}{((g^{r-1} \mod r^2) - 1)/r} \mod r$ [45]. The function is secure if the factorization of $n$ is hard [45]. Size recommendations for $n$ are the same as for RSA [47].

3.1.4 Bao’s Verifiable Encryption

The above mentioned primitives are used in Bao’s Verifiable Encryption (VE) scheme [18] for discrete logarithms. Using VE, a Prover is able to convince a Verifier about the correctness of given encryption of discrete logarithm of some public value. Although the Verifier cannot decrypt, he is convinced that some other entity, which is able to decrypt, will really get the correct discrete logarithm of the given public value after the decryption. The idea is to put the exponent as an input to the Okamoto-Uchiyama function and prove the discrete log equivalence between the original public value and the output of the OU function. After seeing such a proof, the Verifier is convinced without actually seeing the exponent. Later, a third person who knows the OU trapdoor can invert the function and get the exponent (decrypt).

The overview of used cryptographic constructions and their placement in the scheme proposed in Chapter 4 is in Table 3.1.
3.2 Notation

The notation common in cryptographic protocol design is used in this thesis. The same notation is used for both the related work description and the novel scheme design description.

A Discrete Logarithm (DL) commitment \( c \) to a value \( w \) is denoted as \( c = \text{commit}(w) \). For various proofs of knowledge or representation, the efficient notation introduced by Camenisch and Stadler [30] is used. Thus, a protocol depicted in Figure 3.1 can be denoted as \( PK\{w : c = g^w\} \). The proof of knowledge of representation depicted in Figure 3.3 is denoted as \( PK\{w, w' : c = g^w g^{w'}\} \). The proof of discrete log equivalence with respect to different generators \( g_1, g_2 \) depicted in Figure 3.4 is denoted as \( PK\{w : c_1 = g_1^w \land c_2 = g_2^w\} \). The first step of PK protocols is denoted as \( PK_{\text{step1}} \). For example, \( PK_{\text{step1}}\{w : c = g^w\} = (c, \bar{c}) \) in Figure 3.1. Bao’s Verifiable Encryption of \( x \) inside \( c = g^x \mod n \) is denoted as \( \text{VE}\{x : c = g^x\} \). A signature using some PKI by a user U on some data is denoted as \( \text{Sig}_U(\text{data}) \). The symbol “:” means “such that”, “\(|\)“ means “divides”, “\(|a|\)b” is the concatenation of strings \( a \) and \( b \), “\(|x|\)” is the bitlength of \( x \) and “\( x \in_R \{0, 1\}^l \)” is a randomly chosen bitstring of maximum bitlength \( l \). “\( x \in_R \mathbb{Z}_q \)” denotes a randomly chosen integer less than \( q \). “\( \mathbb{Z}_q^* \)” denotes an integer multiplicative group modulo \( q \).
4 NOVEL SCHEME FOR ATTRIBUTE AUTHENTICATION

The cryptographic design of a novel scheme for attribute authentication is provided in this chapter. The proposed attribute authentication scheme supports the basic functionality of anonymously proving the possession of attributes. Additionally, the scheme provides the privacy-enhancing features defined in the thesis objectives chapter. Namely, following features are provided.

- Security
- Anonymity
- Untraceability
- Unlinkability
- Selective disclosure of attributes
- Non-transferability
- Practical, complex revocation

It is shown in the Chapter 2 that it is very difficult to provide some of the above specified features. Especially, the unlinkability and revocation features are very difficult to support. Even the most advanced schemes, like U-Prove and idemix, do not support both these features. The goal of the new scheme proposal is to fix this problem and provide the missing features. Moreover, the scheme design is focused on providing more complex revocation so that service providers are more willing to use the newly designed system in practical scenarios. That is the reason why a scalable revocation mechanism is provided in the proposed scheme. For a practical system, it is also very important to be computationally efficient. Thus, the proposed scheme is designed to be as fast as possible on weak devices such as smart-cards.

The proposed scheme is built on cryptographic constructions defined in Chapter 3. The cryptographic commitments, proofs of knowledge (PK) and verifiable encryption (VE) are frequently used. All constructions are based on the discrete logarithm (DL) assumption. Two domains (multiplicative groups) are used. They are the DSA subgroup modulo prime \( p \) and the Okamoto-Uchiyama subgroup modulo prime product \( n \). Although multiple subgroups are used, the reader might still use the description of cryptographic constructions in the Chapter 3 as a reference. The properties and mechanisms remain the same for all used settings (distinguished by different moduli). The cryptographic background is specified before the actual scheme is described.

Additionally to DSA subgroup, the Okamoto-Uchiyama group is used. The difference is mainly in the used modulus and in the fact that the group order is hidden in OU group. Therefore, the third message of \( \Sigma \)-protocols, the answer, must be sent as a non-reduced integer instead of the reduced answer in \( \mathbb{Z}_q \) sent in the DSA version. The OU group, used in proofs of knowledge and verifiable encryptions, can be easily identified by used modulus \( n \). The full scheme is based on following assumptions.

Assumptions

The Generalized Discrete Logarithm Problem (GDLP) assumption

Based on [38] it is assumed that given a finite cyclic group \( G \) of order \( q \), a generator \( g \) and an element \( \beta \in G \) it is hard to find an integer \( 0 \leq x \leq q - 1 \) such that \( g^x \equiv \beta \).

\(^1\)In our case, the DSA and OU group.
Factorization hardness assumption
Based on [45] it is assumed that it is hard to factor \( n = r^2 s \), where \( r, s \) are large safe primes.

The DSA subgroup, introduced in Chapter 3, is defined by public parameters \( h \) (generator), \( q \) (prime order) and \( p \) (prime modulus). The OU group is defined by public bases \( g_1, g_2, g_3 \) and product modulus \( n \) such that \( n = r^2 s \), where primes \( r, s \) are secret. The commitments, proofs of knowledge, proofs of DL equivalence and verifiable encryptions are used in the full scheme. All primitives are defined in the Chapter 3 which can be used as a cryptographic reference. The difference between the DSA subgroup, from which all examples in Chapter 3 come, and groups with hidden order is mainly in sending the final answer in integers. The communication pattern remains, further information can be found in [33, 18].

4.1 Additional Contribution
In this chapter, an attribute authentication scheme with practical, efficient multi-level revocation is provided. From user’s perspective, the scheme is extremely efficient, providing following features. These features are additional to the privacy-enhancing features defined in Chapter 2 and their purpose is to further enhance privacy and practicality of the scheme.

- **Practical Revocation**
  - **Immediate Revocation**: there is no need to wait for the token lifetime expiration, tokens can be revoked immediately.
  - **Issuer and Verifier Revocation**: revocation is available to both attribute issuers and verifiers. Any of these entities can initiate the revocation process.
  - **Verifier Local Revocation (VLR)**: valid users do not have to update their tokens or download any values after some other users are revoked.
  - **Practical Revocation**: computational complexity of the verification protocol does not depend on the number of revoked users.

- **Scalability**: computational complexity of the verification protocol does not depend on the number of all users.

- **Off-line Verification**: the verification session runs between the user and verifier only. There is no need to contact other parties.

- **Computationally Efficient Verification**: the verification protocol with a revocation check needs only 6 modular exponentiations, 6 multiplications and 4 additions. The complexity of the protocol does not depend on the number of undisclosed attributes\(^2\).

The token revocation is an extremely important feature but in some cases it is not enough to just revoke users from the system. In cases where damage was done, the service providers need a mechanism for learning the identity of attackers to make them responsible. More granularity to revocation is added to the scheme by allowing the revocation of particular privacy enhancing features. The

\(^2\)The protocol independence on undisclosed attributes seems obvious and easy to provide but is currently missing in leading related schemes.
revocation of following features is supported.

- **Token Validity Revocation**: in most practical cases, like token expiration, loss or theft, the token can be revoked without identifying the owner.
- **Unlinkability Revocation**: in non-critical policy breaches, the verifier can inspect user’s past behavior by revoking unlinkability. All past sessions of a particular user can be inspected without releasing his identity.
- **Anonymity Revocation**: in critical policy breaches, it is possible to revoke the anonymity of a user to make him responsible for his acts.

It is acknowledged that these revocation features must be strongly protected against a misuse. That is the reason why the ability to do revocation is spread over more entities. In the system, the issuer, verifier and a third authority must cooperate to revoke any privacy enhancing feature. By such distribution, the probability of misusing the revocation by a single authority is limited. To provide more security (and user trust in the system), the third party can be distributed using multi-party computations. Moreover, the user should have the freedom to choose his own attribute issuer among many commercial subjects, therefore, he does not have to trust a fixed designated revocation authority but rather liberally choose an entity he trusts most.

### 4.2 General Overview

There are 4 types of players in the scheme. They are the User (U), Issuer (I), Verifier (V) and Public Authority (PA). Their roles are described in the following section.

- **User**: a player who wants to get verified using a smart-card with a token issued by I and PA.
- **Issuer**: a player who issues attributes to users and collaborates during revocation. There are more Issuers, the Issuer can be chosen by the User depending on his preferences and trust.
- **Verifier**: a player who verifies Users’ attributes and collaborates during revocation.
- **Public Authority**: a (distributed) entity that issues tokens to Users and collaborates during revocation.

The scheme is based on the interaction among these four entities. The User owns a smart-card with a token construction created using information from both PA and Issuer. The token can be used for attribute authentication, since the Verifier is able to check token’s validity. This is done without the revelation of User’s real identity. But when the User breaks some rules of the Verifier, e.g. destroys some data, the Verifier can ask the Issuer and Public Authority to “open” the token and reveal the identity of the User. The User is then held responsible for his past behavior. There are more levels of User revocation (from token revocation to anonymity revocation) and PA is trusted not to revoke tokens in unjustified cases.

The scheme provides security features which prevent entities from misusing their powers. V alone, I alone or PA alone cannot track or identify a User at all. PA cannot falsely accuse a particular User, but has to be trusted not to misuse its data for accusing random Users during revocation. To lower
such trust, distributed PA can be introduced. Only the joint cooperation of V, I and PA can do the revocation, by sharing their inputs.

The scheme consists of 4 protocols - setup, registration (attribute issuance), verification and optional revocation. User’s goal of the registration is to get data from I and PA necessary to construct the token. The token is then used during the Verification phase to prove User’s membership, attribute ownership or any other authorization given by I’s organization. In case of disputes, loss of a smart-card or system policy breaches, the revocation phase can be introduced. Based on the level of dispute, the User can be removed, traced or identified. The revocation is possible only if I, V and PA cooperate, assuming PA cooperates only in cases where sufficient policy violation logs are given by V’s organization. The communication pattern is depicted in Figure 4.1.

4.3 Cryptographic Design


Setup

The goal of the setup phase is to generate all necessary initial parameters. It is assumed that a public key infrastructure and valid keypairs for the User and Issuer exist. They will be used during registration. The security parameters are \( k, l, m \) (\( k \) is the length of the challenge/hash function used, \( l \) relates to the length of Users’ secrets, and \( m \) is the verification error parameter). The Issuer generates the DSA subgroup \( G \) defined by a large prime modulus \( p \), generator \( h \) of prime order \( q : |q| = l \) and \( q|p − 1 \) . The Public Authority generates groups \( G_1, G_3 \) for the Okamoto-Uchiyama Trapdoor One-Way Function. \( G_1, G_3 \) are defined by the modulus \( n = r^2s \) with \( r, s \) large primes (\( |r| ≥ 350, |r| > 2l, |n| ≥ 1024 \)), generators \( g_1, g_3 \in_R \mathbb{Z}_n \) of order \( \text{ord}(g_1 \text{ mod } r^2) = \text{ord}(g_3 \text{ mod } r^2) = r(r − 1) \) in \( \mathbb{Z}_{r^2}^* \) and \( \text{ord}(g_1) = \text{ord}(g_3) = rr's' \) in \( \mathbb{Z}_n^* \). PA also randomly chooses its secrets \( S_1, S_2 : |S_1| = l, |S_2^{-1} \text{ mod } \phi(n)| = l \) and \( \text{GCD}(S_2, \phi(n)) = 1 \). Finally, PA computes a token \( A = g_1^{S_1} \mod n \) (public, common for all Users) and a value \( g_2 = g_1^{S_2} \mod n \). There might be more types of tokens (different \( A_i \)'s) related to different attributes Users want to prove. In that case, each
unique $A_i$ represents one attribute, e.g. nationality, driving permission or legal voting age. These attributes can be aggregated together by multiplying mod $n$. In the rest of the paper, only one $A$ representing a general group membership is considered for simplicity.

The values $q, p, h, n, g_1, g_2, g_3, A$ are made public, while $r, s, S_1, S_2$ are securely stored at PA or secretly shared in PA’s distributed environment.

**Registration Protocol**

We assume a secure, authenticated (non-anonymous) channel between the User and the Issuer during the registration phase. The User randomly chooses a secret value $w$ shorter than $l$ bits, makes a signed commitment $c_I$ to it and proves the knowledge of $\log_w c_I$ to the Issuer. The User binds his identity to the commitment by the signature. Therefore, the key is computed jointly so that PA does not learn $(w, w')$ but $A$ does. The User’s smart-card computes a credential seed $A' = g_1^w g_2'^{w'}$ mod $n$ where $w', w''$ are User’s random contributions to the key. User sends $A', c_I$ and I’s signature to PA and prove that the key part $w$ is present in both $A'$ and $c_I$. This can be proven in zero-knowledge using $PK\{w : c_I = h^w \text{ mod } p \land A' = g_1^w g_2'^{w'} \text{ mod } n\}$. PA verifies the proof, check if $c_I$ is not used by any other user and answers with his key contribution $w'_{PA}$. The $w'_{PA}$ is computed as $w'_{PA} = ((A * A'^{-1}) \text{ mod } n)^{-1} \text{ mod } r^2 - 1)$ mod $r$ so that $A = g_1^w g_2'^{w'} g_2^{w'_{PA}}$ mod $n$. The User’s smart-card sets $w' = w'_r + w'_{PA}$ and stores the credential $(A, (w, w'))$.

By using the described registration protocol, all Users share the same $A$ for a particular attribute but use different keys. Users are stuck to their keys, because without $S_1, S_2, \phi(n)$ they are unable to compute different valid keys. The registration protocol is depicted in Figure 4.2.

**Verification Protocol**

The verification protocol runs over an anonymous channel (using either anonymous routing or smart-cards). The User proves the knowledge of representation of $A$ using $(w, w')$. By doing so, he remains completely anonymous, since the proof of representation can be a Zero-Knowledge protocol. To get the ability to revoke invalid Users, a subprotocol must be added. User cannot give a commitment to his $w$, since this would make his sessions linkable. Therefore, he creates a verifiable encryption of the randomness used in the first step of the proof of representation protocol (denoted as $PK_{\text{step1}}$).

---

3. A represents a unique attribute. A public list of available attributes and their assigned values of $A$ should be maintained by PA.

4. Due to security reasons, $(w, w')$ are never accessible to users directly but are stored in smart-card’s protected memory. Keys are never extracted, all operations involving them are done on the smart-card.
\[ w \in R \{0, 1\}^{l-1} \]
\[ c_I = \text{commit}(w) = h^w \mod p \]

\[ PK\{w : c_I = h^w\} \rightarrow \text{Sig}_U(c_I) \]
\[ \text{Store } \{c_I, \text{Sig}_U(c_I)\} \leftarrow \text{Sig}_I(c_I) \]

\[ w'_U \in R \{0, 1\}^l \]
\[ A' = g_1^{w_1} g_2^{w_1'} \mod n \]

\[ A', c_I, \text{Sig}_I(c_I), PK\{w : C_I = h^w \wedge A' = g_1^w g_2^{w'}\} \]
\[ w'_{PA} : A = g_1^w g_2^{w_1} g_2^{w_2} \mod n \rightarrow \]
\[ w' = w'_U + w'_{PA} \]

Credentialed: \((A = g_1^w g_2^{w'} \mod n, (w, w'))\)

Fig. 4.2: Registration Protocol

\[ \text{Token} = A = g_1^w g_2^{w'} \mod n \]

\[ PK\{w, w' : A = g_1^w g_2^{w'}\} \rightarrow \text{VE}\{r_1 : PK_{step1} = g_1^r g_2^r\} \]

Fig. 4.3: Simplified verification.

Such encryption will be always different, thus unlinkable. Nevertheless, if decrypted by PA, it can be used to trace a User or even to identify him. The User has to prove that the randomness in VE is the same as the one used in the proof of representation. PA is not involved during the verification at all, therefore the verification, including the revocation check, can run off-line. The protocol is depicted in Figure 4.3. A detailed description of the verification protocol with all the checks needed to support revocation is given in Figure 4.4.

Revocation Protocol

The revocation protocol is the fundamental part of the system, assuming a good balance between anonymity and accountability is required. Practical credential schemes [29, 39] do not implement scalable revocation and theoretical proposals [26] are not efficient enough to be implementable. In the proposed scheme, there is no way of identifying or tracing a User unless more entities cooperate. The verification protocol transcript must be given to (distributed) PA entity with an evidence of a breach of a policy. PA can then decide what type of revocation will be applied: tracing, revocation or identity revelation. The first two types can be done by PA, but in the case of identity revelation, the real identity is readable only to the Issuer (previously chosen by the User), so no unnecessary data about Users ever leak, even if Users break Verifier’s policies.
Tracing (Unlinkability Revocation): the Verifier can give two transcripts of different verification protocols to PA. PA, using its secret factorization of $n$, can decrypt the randomness inside $V^5$, thus it can get User’s unique $w^6$ (but not the real identity!), so it can decide whether two verification protocols were conducted by the same User.

Revocation (Token Revocation): based on a verification transcript, PA can get $w$ like in the Tracing case. To prevent identity revelation, PA can release a commitment to $w$ in the form of $\text{rev} = g^w \mod n$. Using $\text{rev}$, the Verifier can always check whether the User has been revoked or not without de-anonymizing the User. The revocation check is described in Figure 4.4.

Identity Revelation (Anonymity Revocation): in the worst cases of rule breaking, PA can reveal $w$ to the Issuer, who can compute an adequate commitment $c_I = h^w \mod p$ and charge the User who signed that commitment during the registration.

The above described Revocation Protocol provides full anonymity to honest users. Malicious users, whose communication is submitted to one of three types of revocation, can be revealed, because their secret keys can be disclosed by PA. In our design, PA is trusted not to disclose User’s secrets if insufficient policy violation proofs are submitted. It is the next goal of following research to design a scheme were such a trust in PA is removed. In the time of writing this thesis, a solution based on disclosing commitments of keys is being developed.

In this chapter, the novel scheme for anonymous attribute authentication was presented. Using the scheme, a User can anonymously convince a Verifier about the possession of an attribute, typically about the authorization to use services. By staying anonymous and having the control over all released data, the users can protect their privacy during the verification process. The full scheme supports all required features, so far supported only individually, in one practical solution, which is even easily implementable on smart-cards. Additionally to being computationally efficient, the scheme can be proven secure\(^7\), since it is based on strong cryptographic primitives, mostly proofs of knowledge or representation, Okamoto-Uchiyama OWF and DL commitments.

Additionally, features unavailable before are included, mainly scalable and implementable revocation with malicious user identity revelation. These new features make the scheme acceptable by service providers, without whose support any scheme is without a chance for being widely accepted.

\(^5\)With respect to Figure 4.4, randomness is obtained as $r_1 = (r_2^{-1}) \mod r^2 - 1)/r \mod r$

\(^6\)With respect to Figure 4.4, User’s secret can be obtained as $w = ((r_1 - z_1)c_2^{-1}) \mod r$

\(^7\)Security depends also on underlying tools, like the anonymous routing protocol (TOR) and the smart-card.
Fig. 4.4: Verification protocol (optimized).
The main objective of this thesis is to bring more privacy and digital identity protection to user authentication. This intention is motivated by the increasing number of new types of services and technologies which bring new threats to our privacy and digital identity. Technologies like electronic IDs, cloud services or ubiquitous smart-cards intensively work with our personal data although we cannot be sure if the protection is adequately high. Thus, a novel cryptographic scheme assuring personal data protection is proposed in this thesis. The scheme directly reacts on official U.S. and EU requests [44, 42] for new authentication services with the support of user-centric, attribute-based approach to authentication. Besides privacy-enhancing features, the proposed scheme provides also higher security because it is built on strong cryptographic constructions with provable features.

The proposed scheme is based on a new concept called attribute authentication. In this concept, a user can give proofs not only about his identity but also about the possession of attributes. The attributes represent any personal data, e.g. age, citizenship or nationality. With the increasing number of services, where identity disclosure is not necessary but the verification of some attribute is required, the user can anonymously prove that he is a justified holder of that attribute. By using attribute authentication, honest users can give proofs about their age, nationality or any authorization while staying completely anonymous. To protect assets of verifiers, the scheme allows the de-anonymization of malicious users so that they can be held responsible for their behavior. By allowing attribute authentication, the users can precisely manage what personal data are being released and are assured that the verifier is unable to identify, trace and profile their behavior.

An extensive analysis of current state in the field of cryptographic privacy protection is provided in this thesis. Due to missing support of privacy protection in classical protocols like RADIUS, EAP or Diameter, more advanced cryptosystems have been analyzed. The most promising technologies, U-Prove and idemix, were examined. Major weaknesses were found in these systems, namely missing unlinkability of verification sessions and weak revocation.

Due to the absence of these important features, a new scheme supporting all privacy-enhancing features including unlinkability and advanced revocation was designed. Particularly, it is the revocation feature, which virtually prevented existing systems from being commercially enrolled. Revocation was the main topic of many research papers from the last decade but it has been still unresolved, especially on low-performance, off-line devices. Without revocation, it is impossible to revoke invalid, expired, lost or stolen authentication tokens. The scheme presented in this thesis is the first solution which is both providing practical revocation and highly practical on devices like smart-cards. With the proposed scheme, it is now possible to significantly enhance the privacy of users of electronic services. The solution is the first one which provides strong authentication with privacy protection together with the protection of service providers. Moreover, the scheme is practical on devices like off-line smart-cards which allows the application to electronic IDs.

An extensive verification of the proposed scheme is presented in the full version of the thesis. Both theoretical and practical verification is included. Three key properties, soundness, completeness and anonymity of verification phase, are formally proven. The scheme is verified on a mathematical model using Mathematica software. An experimental smart-card implementation of the verification phase is provided. The time of verification is around 30% faster than in related existing implementations. The results have been published at international conferences, awarded prizes (Keymaker
2011, Brno Ph.D. Talent, Fulbright Stipend) and presented at key institutions (University of Minnesota, National Institute of Standards and Technology (NIST), USA, Goethe-University Frankfurt (ABC4Trust Project Leader), DE).

Although the proposed scheme is fully functional, there is a potential for future development. The full implementation has been currently started as a TACR project with a commercial partner. Also, there are several aspects of the scheme which can be cryptographically improved. These aspects are being investigated by the author and foreign partners. The goal is improving the verification and revocation procedure so that user keys are not fully learnt by PA, more attributes are proven more quickly and the trust in key entities like PA is even more lowered. All these three improvements are the core subjects of follow-up research.
BIBLIOGRAPHY

Author’s selected publications


Other’s publications


Jan Hajný

Affiliation: Brno University of Technology, Brno, CZ

Contact: +420608823522, hajny@feec.vutbr.cz

EDUCATION

2008 - 2012 Doctoral program at Brno University of Technology, CZ
Faculty of Electrical Engineering and Communication
Research topic: “Authentication protocols and privacy protection”

2010 - 2011 Department of Mathematics, Department of Computer Science
University of Minnesota, USA

2008 Department of Computer Science (Crypto-Group)
University of Arhus, Denmark

2006 - 2008 Master’s program at Brno University of Technology, CZ
Master’s thesis: “Analysis and design of authentication systems”
Grade average for whole program (1 - excellent 4 - failed): 1.08

2003 - 2006 Bachelor’s program at Brno University of Technology, CZ
Bachelor’s thesis “Methods for LAN protecting”
Grade average for whole program (1 - excellent 4 - failed): 1.37

WORK EXPERIENCE

2008 - Present Computer and communication security lab instructor
2006 - 2007 Network security consultant

LANGUAGE SKILLS

Czech language Native
English language Fluent, certified
German language Intermediate

CERTIFICATIONS

General Graduate Record Examination GRE General Test (780/800 in QR)
Language University of Cambridge, Certificate of Advanced English - C1 Level
ETS TOEFL (Test of English as a Foreign Language) iBT, score 108/120

PRIZES AND HONORS

2011 1. Prize in Keymaker cryptography competition in Ph.D. section
2010 Fulbright stipend for internship at University of Minnesota, USA
2010 Brno Ph.D. Talent prize
2010 AVG prize for security projects
2009 EEICT 2009 - first prize in Ph.D. section “Information Systems”
2008 Master degree with honors
2006 Bachelor degree with honors

SELECTED PROJECTS

2012 - 2014 TA02011260: Cryptographic system for the protection of electronic identity
2012 - 2014 FR-TI3/170: Integration server with cryptographic protection
2009 - 2012 JCM M Brno Ph.D. Talent: Cryptographic identity protection
2012 FRVŠ 823/2012: Innovation of cryptography and computer security courses
2010 FRVŠ 2833/2010: Inclusion of application firewalls to security courses
2010 FEKT-J-10-5: Authentication and authorization in modern computer networks
2008 - 2011 2C08002: Universal and complex authentication and authorization for computer networks
SELECTED INVITED PRESENTATIONS
2012 Authentication and privacy protection at SmartCard Forum
2011 Privacy Protection by Attribute Authentication at NIST
2011 Cryptographic System for Identity Protection at MKB 2011

PUBLICATIONS IN NUMBERS
Impacted journals 3 publications
Journals 6 publications
Int. conferences 7 publications
Czech conferences 7 publications
Total 23 publications

SUPERVISED COURSE LABS
2008, 2009, 2011 Information systems security
2009, 2010 Design, administration and security of computer networks
ABSTRACT

This dissertation thesis deals with the cryptographic constructions for user authentication. Rather than classical authentication protocols which allow only the identity verification, the attribute authentication systems are the main topic of this thesis. The attribute authentication systems allow users to give proofs about the possession of personal attributes. These attributes can represent any personal information, for example age, nationality or birthplace. The attribute ownership can be proven anonymously and with the support of many features for digital identity protection. These features include, e.g., the unlinkability of verification sessions, untraceability, selective disclosure of attributes or efficient revocation. Currently, the attribute authentication systems are considered to be the successors of existing authentication systems by the official strategies of USA (NSTIC) and EU (ENISA). The necessary features are partially provided by existing cryptographic concepts like U-Prove and idemix. But at this moment, there is no system providing all privacy-enhancing features which is implementable on computationally restricted devices like smart-cards. Among all weaknesses of existing systems, the missing unlinkability of verification sessions and the absence of practical revocation are the most critical ones. Without these features, it is currently impossible to invalidate expired users, lost or stolen authentication cards and cards of malicious users. Therefore, a new cryptographic scheme is proposed in this thesis to fix the weaknesses of existing schemes. The resulting scheme, which is based on established primitives like \( \Sigma \)-protocols for proofs of knowledge, cryptographic commitments and verifiable encryption, supports all privacy-enhancing features. At the same time, the scheme is easily implementable on smart-cards. This thesis includes the full cryptographic specification, the formal verification of key properties, the mathematical model for functional verification in Mathematica software and the experimental implementation on .NET smart-cards. Although the scheme supports all privacy-enhancing features which are missing in related work, the computational complexity is the same or lower, thus the time of verification is shorter than in existing systems. With all these features and properties, the resulting scheme can significantly improve the privacy of users during their verification, especially when used in electronic ID systems, access systems or Internet services.