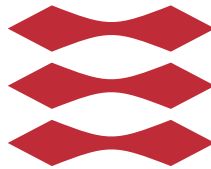


Non-identifying Data Management Systems

Dheeraj Kumar Bansal

DTU



Kongens Lyngby 2015

Technical University of Denmark
Department of Applied Mathematics and Computer Science
Richard Petersens Plads, building 324,
2800 Kongens Lyngby, Denmark
Phone +45 4525 3031
compute@compute.dtu.dk
www.compute.dtu.dk

Summary (English)

The goal of the thesis is to ...

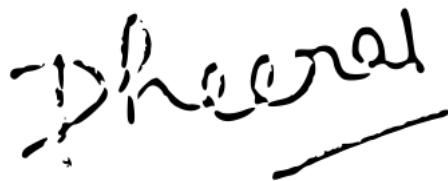
Preface

This thesis was prepared at DTU Compute in fulfilment of the requirements for acquiring an M.Sc. in Engineering.

The thesis deals with ...

The thesis consists of ...

Lyngby, 26-June-2015

A handwritten signature in black ink, appearing to read 'Dheeraj', with a long horizontal stroke underneath.

Dheeraj Kumar Bansal

Acknowledgements

I would like to thank my....

Contents

Summary (English)	i
Preface	iii
Acknowledgements	v
List of Figures	xi
1 Problem Statement and Background	1
1.1 Introduction	1
1.2 Background	1
1.3 Definitions	2
1.3.1 Privacy	2
1.3.2 Anonymity	2
1.3.3 Pseudonym	2
1.3.4 User	2
1.3.5 Bank	3
1.3.6 Third party	3
1.3.7 Service	3
1.3.8 Unlinkability	3
1.3.9 Revocation	3
1.3.10 Partial Information Disclosure	3
1.3.11 Conditional Anonymity Removal	4
1.3.12 Legal Requirement	4
1.4 Case Study	4
1.4.1 Private Customers	5
1.4.2 Corporate Customers	5
1.5 Problem	5
1.6 Summary	6

2	State of the art Survey	7
2.1	Introduction	7
2.2	Technologies	7
2.3	Secure Multi-Party Computing	8
2.3.1	Homomorphic Encryption	8
2.3.2	Group Signature	9
2.4	Escrow Technologies	9
2.4.1	Secure Logging	9
2.4.2	Threshold Cryptography	9
2.5	Identity Management Systems	10
2.5.1	OpenID	10
2.6	Zero Knowledge Technologies	11
2.6.1	IDEMIX	11
2.6.2	U-Prove	15
2.7	Summary	17
3	Application Scenario	19
3.1	Introduction	19
3.2	Information Flow	19
3.3	Separation of Identities	21
3.4	Current Banking System	25
3.5	Summary	26
4	Proposed System	27
4.1	Introduction	27
4.2	User Identification	27
4.3	Pseudonym System	28
4.4	Properties	29
4.5	User Privacy	29
4.6	Summary	30
5	Simple Prototype	31
5.1	Introduction	31
5.2	Design	31
5.3	Authentication	32
5.4	User Information	33
5.5	Transactions	34
5.6	Summary	38
6	OpenID Based Solution	39
6.1	Introduction	39
6.2	OpenID IMS	39
6.3	System Setup	40
6.3.1	Changes on the Bank Side	40

6.3.2	Information Stored at IMS	40
6.3.3	Changes needed on the User Side	41
6.4	User Creation	41
6.5	User Authentication	41
6.6	ID Escrow	42
6.7	Summary	42
7	IDEMIX Based Solution	43
7.1	Introduction	43
7.2	IDEMIX IMS	43
7.3	System Setup	44
7.3.1	Changes on the Bank Side	44
7.3.2	Information Stored at IMS	44
7.3.3	Changes needed on the User Side	45
7.4	User Creation	45
7.5	User Authentication	45
7.6	ID Escrow	46
7.7	Summary	46
8	Analysis	47
8.1	Introduction	47
8.2	OpenID Based pseudonym System	47
8.3	IDEMIX Based pseudonym System	48
8.4	IDEMIX implementation in the Real World	49
8.4.1	Addition of the New User	49
8.4.2	Addition of a New Customer	51
8.4.3	Technical Requirements	51
8.5	Summary	52
9	Conclusion and Future work	53
9.1	Introduction	53
A	Appendix A	55
	Bibliography	57

List of Figures

1.1	Identities in the system	4
2.1	Technology Overview	8
2.2	IDEMIX Roles	12
2.3	IDEMIX Credential	13
2.4	IDEMIX Presentation Token	14
3.1	Example of information maintained for each relationship	20
3.2	Current Banking System	25
4.1	Pseudonym Banking System	28
5.1	Prototype Design	31
5.2	User Login Page	32
5.3	Authenticated User	32
5.4	User Information - Session 1	33
5.5	User Information - Session 2	34
5.6	New Transactions	35
5.7	Account Transactions	36
5.8	Download Transactions Option	36
5.9	Downloaded Transactions File	37
6.1	Pseudonym System with OpenID IMS	40
7.1	Pseudonym System with IDEMIX IMS	44
8.1	Pseudonym System with OpenID IMS	47
8.2	Pseudonym System with IDEMIX IMS	48
8.3	IDEMIX Credential issuance for a new user	50

8.4 Final IDEMIX Credential from Policy Credential 50

CHAPTER 1

Problem Statement and Background

1.1 Introduction

This chapter introduce the topic of the thesis. It defines the problem statement and also give a brief background about it.

1.2 Background

Administrative data systems currently rely on the Central Personal Register Id (CPR Nr.) to link customer data with a real world identity. This means that almost all data managed by the institutions must be classified as personal identifiable information and therefore managed according to strict confidentiality requirements as well as integrity and availability requirements. The purpose of this thesis project is to examine ways to decouple transaction data from the underlying identities, so that the data used in the institution's day to day operations are decoupled from the underlying customer identity. This limits the confidentiality requirements for the data and the vulnerability to insider threats, such as the recent leak of celebrity data from NETS to the magazine "Se & Hør".

It must, however, be possible to link customer data to real world identities when reporting financial data to the Tax authorities or in connection with suspicions about criminal activity, e.g. fraud, insider trading, whitewashing, etc. Austria is already considering similar approaches in the public health care system, where the health records of the citizens are saved under pseudonyms, which are mapped one-to-one to the single citizen.

1.3 Definitions

Here we will give some of the definitions used in the system

1.3.1 Privacy

Privacy is the ability of an individual to control the distribution of information about himself. An individual should be able to choose which information about him should remain secret and which information can be revealed.

1.3.2 Anonymity

Anonymity refers to ability of a user to not give any information about him at all to the system. An anonymous system doesn't have any identity of the user.

1.3.3 Pseudonym

Pseudonym is a name given to the user in pseudonymous systems. This name is given to hide the real identity of the user from the system. The system only knows the user by his pseudonym.

1.3.4 User

User is the end user of the system. Its the person who will go online and get the services.

1.3.5 Bank

Bank is the financial institution which provides online financial services to the user.

1.3.6 Third party

A third party or trusted third party is the entity which is neither bank or the user in the system. A third party provides different services to banks or users and hence reducing the burden on them to setup all the infrastructure by themselves.

1.3.7 Service

Service is something that is provided to the user online by a system. It may include ability to login, check his account balance, Upload pictures, share spreadsheets etc.

1.3.8 Unlinkability

Unlinkability is the privacy property where its not possible to link 2 different entities to each other even though they are the same.

e.g. not be able to link 2 different sessions by the same user in the system.

1.3.9 Revocation

Revocation is the property where a user credential is revoked by user or some other authority. After revocation, this credential cannot be used for anything.

1.3.10 Partial Information Disclosure

Its the ability of a user to only disclose some partial information about himself to the system. e.g. a user might just want to disclose his last name to the

system but not his full name.

1.3.11 Legal Requirement

A legal requirement is something that is required by the law. e.g. it may be required by the law for the bank to log all the customer data. Also sometimes in case of suspicious transactions bank maybe required legally to give the user identity to the relevant authorities.

1.3.12 Conditional Anonymity Removal

Its the ability of the system to remove anonymity of the user if some conditions that were set before are met. This is mainly used for escrow purposes.

1.4 Case Study

Nykredit is a major financial institution in Denmark providing different services, such as mortgages, retail banking, investment banking etc. They also are part of a big group of companies, which includes other financial institutions providing similar services. These financial institutions basically provide Nykredit services as their own services to the customers. Nykredit has mainly 2 types of customers:

- Private Customers
- Corporate Customers

1.4.1 Private Customers

Private customers are the individual customers who access Nykredit services on their own. Usually there is a single person accessing the services of the bank. These customers are usually people who get a personal bank account with Nykredit.

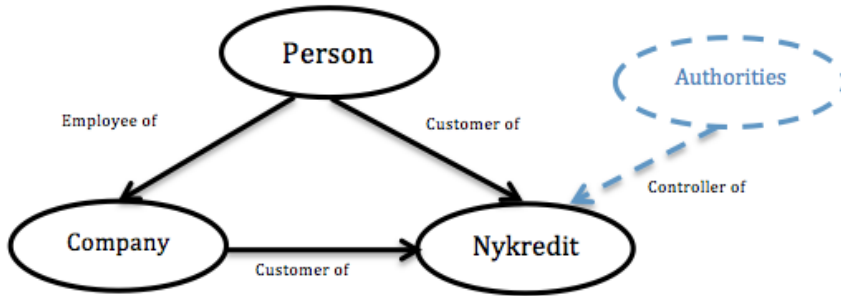


Figure 1.1: Identities in the system

1.4.2 Corporate Customers

Corporate customers are either companies who are customers of Nykredit or other financial institutions which provide Nykredit services to their own private customers. Usually there are many people who access Nykredit services on behalf of the corporate customer.

1.5 Problem

We consider the case of a person, who may either be a private customer of Nykredit, or an employee of a company who is corporate customer of Nykredit. In this case, the person may also be responsible for managing the accounts of his employer with Nykredit.

Nykredit wants to setup an identity management system so that there is no need for the individual to disclose his personal identity to Nykredit to access the account on behalf of the company.

Nykredit, however, also have to comply with relevant legislation (KYC, AML, “Hvidvaskningsloven”), e.g. in case the authorities (Tax, Police, etc.) find some suspicious transactions. Nykredit needs to provide the identity of the person responsible for these transactions. This means that it is required that Nykredit, in case of a legal request, is able to identify the individual employee from the institution, who is accessing the account on the corporate customer’s behalf.

So the main goal of the system is:

- Nykredit should not learn identity of the individual person accessing the services on behalf of corporate customer.
- For complying with legislation Nykredit should be able to map real identity of the individual person with the transaction in case its required by the law.

The project will perform an initial analysis of a single business process from administrative data management, with respect to identifying the need to bind authenticated identities to actions at the different steps of the process; this analysis will be presented to stakeholders from the specific administrative domain. Based on the initial analysis of the selected business process, the project will develop a full identity model for the chosen business process with anonymisation and pseudonymisation of actors whenever possible. The feasibility of the proposed model will be evaluated through a prototype that implements the model using standard components from identity management infrastructures whenever possible.

1.6 Summary

Companies do not want to disclose the personal identity of their employees to Nykredit, but they still need the ability to access all services online. Managing all identities, while maintaining privacy, is not easy and provides different challenges. We have to design a system, which fulfill the entire privacy requirement and still enables Nykredit to provide its services to its customers and meet the regulatory requirements of the authorities without any problem.

CHAPTER 2

State of the art Survey

2.1 Introduction

This chapter introduce all the different technologies currently available which deal with anonymity and privacy. Some of the technologies are already commercially available and being used in the industry, while some are still in research phase. We will give a brief introduction for all of them and brief idea how they can be used.

2.2 Technologies

1. Secure Multi-Party Computing
2. Escrow Technologies
3. Identity Management Systems
4. Zero Knowledge Technologies

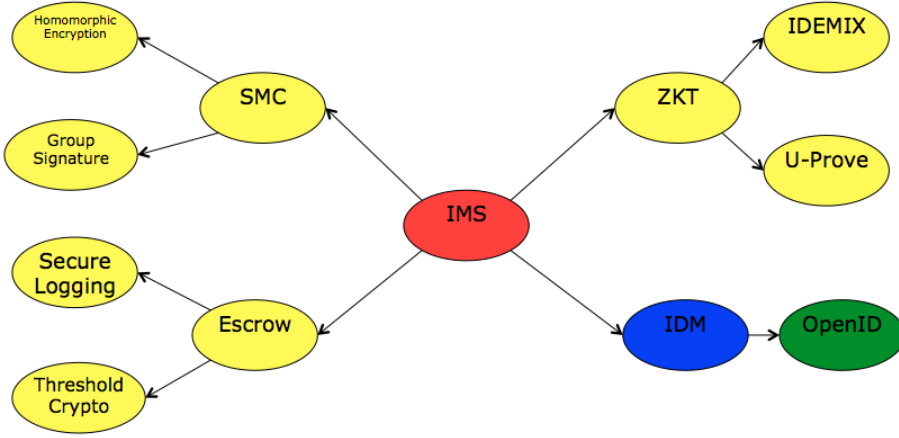


Figure 2.1: Technology Overview

2.3 Secure Multi-Party Computing

These technologies are the ones which involve multiple parties to do computations. It is a subfield of cryptography which involves multiple parties getting an input and compute a joint function on them while keeping these inputs private. We basically looked at 2 technologies of interest here:

2.3.1 Homomorphic Encryption

Homomorphic encryption is a type of encryption where certain arithmetic operations can be performed on the ciphertext such that when the resultant ciphertext is decrypted, the decrypted text is same as the operations were performed on the plaintext. This is a new field of cryptography and is very useful where we need some parties to perform such operations without revealing the underlying data to those parties. Homomorphic encryption is also useful for chaining of different services without exposing data to any of those services.

2.3.2 Group Signature

A group signature is a scheme which allows a member of the group to sign the message on behalf of the group but anonymously. To outsiders the message has been signed by someone from the group but the exact identity of the person is now known. Also if the same member signs 2 different messages its not possible to know if the message is signed by the same member. There is a notion of group manager in these scheme. Group manager is someone who manages the membership to the group. He can add/remove members from the group, find out who actually signed the message from the group. This scheme is useful where only thing that needs to be validated is that a certain person is part of the group, but his real identity is not required.

2.4 Escrow Technologies

Escrow technologies are the ones which are helpful in escrow purposes i.e. getting real data/identity later on in time from encrypted data if needed. We look at following 2 technologies.

2.4.1 Secure Logging

Secure logging is the process of saving the data in a secure manner. As saved data is really crucial and vulnerable to the attacks. We need to make sure that data is saved securely and its integrity is protected. This can be done in several ways. One way is to encrypt all the logs while storing them so that even if someone get hold of the logs, they can't use them without having access to the decryption key. Other way is to store logs at a third party after encrypting them. For escrow purposes, these logs can be decrypted later on with the decryption key.

2.4.2 Threshold Cryptography

Threshold cryptography is a field of public key cryptography where in order to decrypt an encrypted message, several parties must cooperate in the decryption. This message is encrypted using a public key and the corresponding private key is shared among different parties who will participate in the decryption process i.e. multiple parties hold the private key for a single public key. There is a term

called threshold, and if there are n parties who share the private key and at least t parties which are required to decrypt the message such system is called (t,n) threshold cryptosystem. Threshold cryptosystem is useful in escrow purposes where a minimum number of parties can be defined to decrypt the ciphertext in order to get the plaintext.

2.5 Identity Management Systems

These are traditional identity management systems. For our purposes we look at OpenID system.

2.5.1 OpenID

OpenID is an open and decentralized protocol, which can be used to authenticate to different co-operating sites with the use of a third party service. It has notion of a *relying party* and *OpenID identity provider*.

- **OpenID Identity Provider** OpenID identity provider is the service, which actually provides authentication services. End user registers at OpenID identity provider to get his OpenID identity.
- **Relying Party** Relying party is the website which user wants to authenticate to and which rely on the OpenID identity provider to provide authentication.

In addition to this an extension called *OpenID attribute exchange* helps facilitate the transfer of user attributes from identity provider to the relying party

- User goes to the relying party service page
- Service page presents different OpenID providers to login to the service
- User chose the provider he has registered his openID with
- Relying party redirects the user to the OpenID provider url so that user can authenticate
- User can be authenticated by the method provided by OpenID provider

- OpenID provider asks user the permission to share the attributes with the relying party
- After user consent, user is redirected to the relying party website with user credentials
- Relying party can verify the credentials and then login the user to the service

2.6 Zero Knowledge Technologies

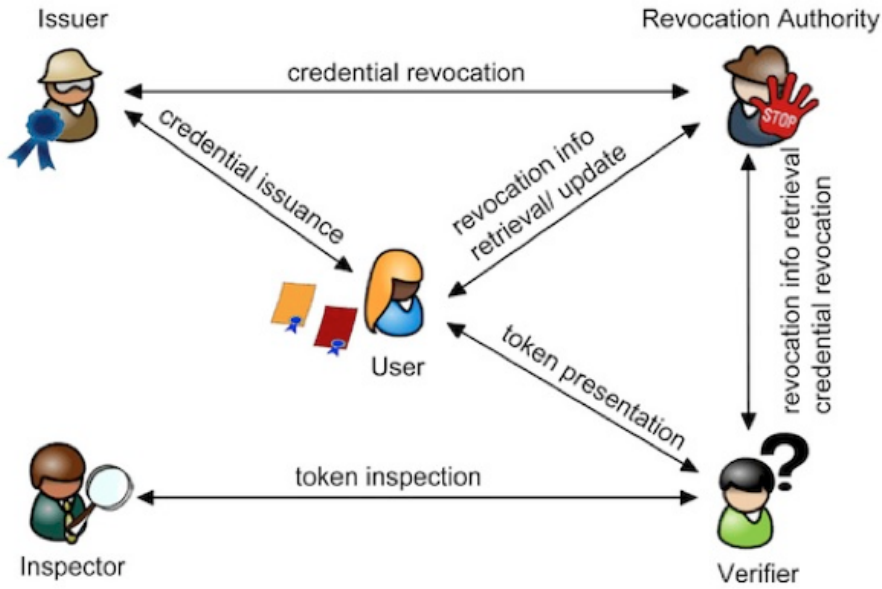
These are the technologies which use the concept of zero knowledge i.e. proving knowledge about something without divulging the information. For our purpose we focus on mainly 2 technologies – IDEMIX and U-Prove, which are based on the concept of zero knowledge and verifiable encryption.. They both have a lot in common and have been studied a lot in industry.

2.6.1 IDEMIX

U-Prove is a digital credential technology by IBM. it relies on anonymous credentials known as IDEMIX tokens. It is based on Camenisch-Lysyanskaya (CL) signature scheme which provides efficient zero-knowledge proofs. IDEMIX have different entities in the system

- **User** User is basically the entity, which is proving his identity in the system.
- **Verifier** Verifier is the entity, which verifies the identity of the prover.
- **Issuer** Issuer is the entity, which issues the credentials to the prover to prove his identity
- **Inspector** Inspector is the entity, which in case of discrepancy or legal requirement , can actually come and get the real identity of the prover.

For IDEMIX we need computing devices which work on behalf of each entity in the system.



Source: <https://github.com/p2abcengine/p2abcengine/wiki/Concepts-and-features>

Figure 2.2: IDEMIX Roles

2.6.1.1 IDEMIX Credential

An Idemix credential is CL Signature by issuer on user's private key and on attribute values. A user have independent public keys or pseudonyms for the same private key. These pseudonyms are IDEMIX tokens which are then used by the user to prove his identity to the different verifiers. IDEMIX has been studied a lot and many EU projects on anonymous credentials are based on it e.g. FutureID, ABC4Trust etc.

2.6.1.2 Issuance

The first step is the credential issuance. It involves following steps:

- User sends credential request to the issuer
- Issuer presents the issuance policy specifying
Which attributes to present

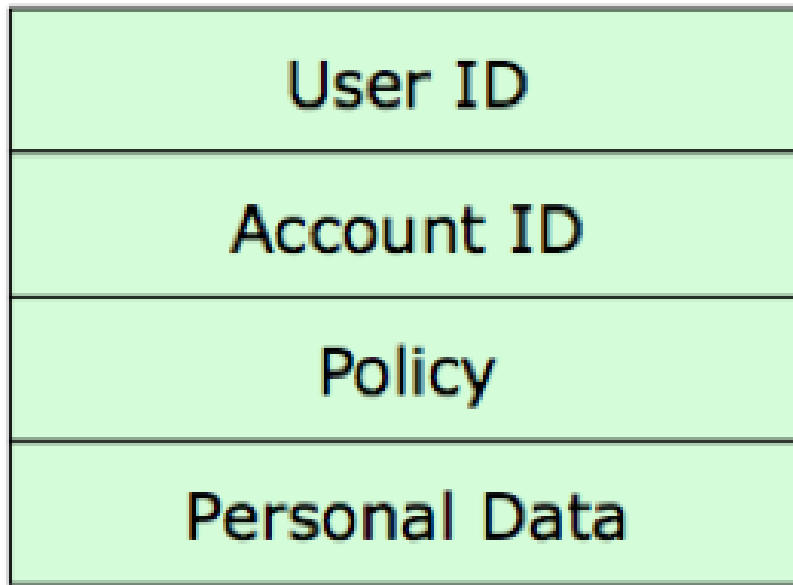


Figure 2.3: IDEMIX Credential

Which pseudonym/existing credentials to present

- Issuer also present a credential template specifying
 - Which attributes of the new credentials will be generated at random
 - Or carried over from existing credential or pseudonym
- User then present the issuance token satisfying the issuance policy
- Then multi-round cryptographic protocol ensues at end of which user get the IDEMIX credential

2.6.1.3 Presentation

Next step is to present the IDEMIX token for authentication to the verifier. It consists of following steps:

- User get the presentation policy from the verifier which specifies

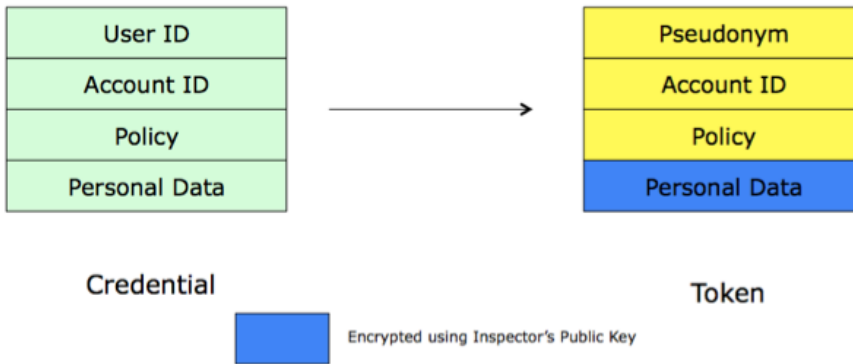


Figure 2.4: IDEMIX Presentation Token

Which credentials user must present

What information user should reveal from these credential

- User generate a presentation token in accordance with the presentation policy revealing only the attributes necessary
- User present this presentation token to the verifier
- Verifier can then verify the attributes

2.6.1.4 ID Escrow

IDEMIX provides the ID escrow ability in case it is required. Following steps need to be followed for ID escrow purposes in IDEMIX:

- The presentation policy can have following optional specifications for the purpose of ID Escrow
 - Public keys of inspectors
 - Attributes values to be encrypted using the keys
 - Inspection conditions under which these attributes can be revealed
- User can prove that he has put these values in the presentation token with verifiable encryption to the verifier

- Once the token is presented, the inspection conditions are fixed and cannot be changed
- In case of some discrepancy or legal requirement, an inspector can come and get the identity of the user from the token

2.6.2 U-Prove

U-Prove is a digital credential technology by Microsoft. It relies on anonymous credentials known as U-Prove tokens. It provides users a way to minimally disclose their personal information while interacting with different online services. U-prove have different entities in the system

- **Prover** Prover is basically the entity, which is proving his identity in the system.
- **Verifier** Verifier is the entity, which verifies the identity of the prover.
- **Issuer** Issuer is the entity, which issues the credentials to the prover to prove his identity
- **Auditor** Auditor is the entity, which in case of discrepancy or legal requirement, can actually come and get the real identity of the prover.

For U-Prove we need computing devices which work on behalf of each entity in the system.

2.6.2.1 U-Prove Token

U-Prove token is basically cryptographically protected information of any kind e.g. attributes. These are issued by issuer to the prover by issuance protocol. These tokens are then presented by prover to the verifier. Issuance and presentation of U-Prove tokens is unlinkable.

2.6.2.2 Issuance

The first step is the credential issuance. It involves following steps:

- Prover invoke U-Prove issuance protocol

- Prover provides the attributes to be encoded

Using *Collaborative Issuance* property user can make sure that issuer doesn't actually know the attributes

- Then multi-round cryptographic protocol ensues at end of which user gets the U-Prove token from the issuer

2.6.2.3 Presentation

Next step is to present the U-Prove token for authentication to the verifier. It consists of following steps:

- Prover invokes the U-Prove presentation protocol
- User generates a presentation token in accordance with the presentation policy revealing only the attributes necessary
- User presents this presentation token to the verifier
- Verifier can then verify the attributes

It must be noted that a revocation check can be added if needed before verifying the token.

2.6.2.4 ID Escrow

ID Escrow in U-Prove is actually an extension to existing U-Prove technology. It uses a type of ElGamal encryption which is verifiable.

- During the presentation protocol, prover proves that his ID is encrypted in the token by the use of verifiable encryption technology
- De-anonymization cannot be done by verifier or issuer
- A special entity called Auditor is responsible for de-anonymization in case of some discrepancy or legal requirement
- Threshold cryptography can be used in case of auditors and key can be split among multiple auditors.

2.7 Summary

All these different technologies provide different level of anonymization in the system. Some of them are easy to integrate in existing technology, while some are still not mature enough. For our purposes from now on we focus on mainly 2 technologies:

- OpenID
- IDEMIX

CHAPTER 3

Application Scenario

3.1 Introduction

This chapter will discuss the information flow of the current system. We will present our understating of the current banking system. we will also give different types of data that exist in the current system and the different operations that are necessary to support the system.

3.2 Information Flow

A person has different information associated with him.

To Nykredit it can be:

- Personal

This Personal ID can be one of the following identifiers:

- External ID (e.g. NEMid)
- Internal ID (e.g. login credentials of the bank)

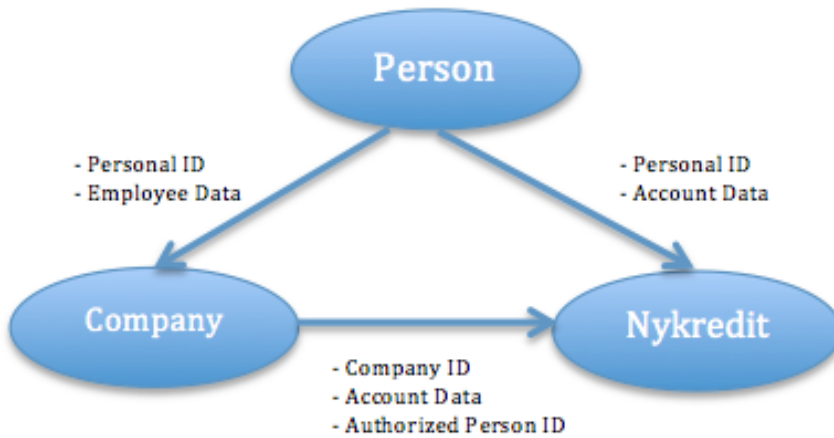


Figure 3.1: Example of information maintained for each relationship

- Account Data

To Company it can be:

- Personal ID
- Employee Data

A company has following information that it might share with Nykredit:

- Company ID
- Account Data
- Authorized Person ID

This Authorized Person ID is the identifier that is used by the authorized person on behalf of the company. This can be stored in some database where it is matched to Personal ID of the person.

The Authorized Person ID can either be same identifier as the Personal ID, the Company ID or a different identifier that can be authenticated.

In case the Personal ID is used as the Authorized Person ID, it gives Nykredit some additional capabilities:

- Nykredit can use this information to recruit new customers. If the authorized person is not a customer before, Nykredit can use this info to contact them.
- If the person is already a customer, then Nykredit can use this information to provide additional services to him on his personal account, so as to influence the authorized person for his decisions regarding the company account (similar to the way airlines reward frequent flyers).
- As Nykredit already knows about company accounts, the performance of the company might influence their decision regarding the private account of the employee (e.g. it may be difficult for a person to take out a mortgage if Nykredit knows that the company they work for is in financial difficulties).
- In the case where the customer is accessing Nykredit services on behalf of a smaller financial institution, the capability of matching the authorized ID to Personal ID gives a chance to Nykredit to recruit this customer away from the smaller financial institutions.

While good corporate governance at Nykredit will prevent these issues, it is desired to completely and demonstratively remove the link between Personal ID and Authorized Person ID. This, however, creates difficulties with respect to regulatory requirements for accountability at Nykredit and KYC, AML, "Hvidvaskningsloven, etc., so there must be some way to map the identifier used in a financial transaction to a real person, i.e. map a given Authorized Person ID to a Personal ID.

3.3 Separation of Identities

A way to remove the link between Personal ID and Authorized Person ID is to use separate IDs and to maintain a database which links the Authorized Person ID back to the Personal ID. This database can be protected in different ways, so that the information can be used only in case legal authorities need to link the two IDs.

This database can be maintained at 3 places :

- Company
- Nykredit
- Trusted 3rd party

3.3.0.5 Database on Company Side

Advantages

- Nykredit doesn't have to invest extra in IT infrastructure

It is expensive to maintain the entire IT infrastructure by Nykredit, so it is easier and cheaper for Nykredit to let the company maintain the database.

- Nykredit can easily prove that it cannot link different Identities

Nykredit does not have access to the database, so it can easily be proved that Nykredit cannot link the different identities.

- Company maintain their own private data

Companies can be sure that Nykredit does not have access to the personal data of their employees

Disadvantages

- There is no way to retrieve data if the company stops existing

In this case the entire mapping database may be lost.

- Authorities have to go to the Company to get the data

Nykredit does not have access to the database, so the authorities have to go to Nykredit first to obtain the Authorized Person ID and then to the individual companies next to get the Personal ID.

- Company might tamper with the database

In the case of a rogue employee at the company, which is exactly the case that the legislation is intended to identify, this employee will have

easy access to this database and hence the ability to tamper with the database and remove the authorities ability to identify him.

- It might prove too difficult for new customers to fulfill all the technical requirements

Larger companies can have their own IT infrastructure, but for smaller companies it might prove to be a difficult task to become a new Nykredit customer if they have to invest extra in IT infrastructure just for this purpose.

3.3.0.6 Database on Nykredit Side

Advantages

- In case the company stops existing, the data can still be retrieved
The database is always with Nykredit, so if some company stops existing, it can still be accessed.
- Authorities have a single place to obtain all the data
Authorities do not have to go to individual companies to get the relevant data as everything is at one place.
- Company cannot tamper with database
As companies have no direct access to the database, they cannot tamper with it.

Disadvantages

- Nykredit has to invest extra in IT infrastructure

Nykredit has to invest extra to keep this system in place.

- Company does not have control over their own private data

The database is on Nykredit side, so companies have to store the data there and hence they do not have control over their own private data.

- It is difficult for Nykredit to prove that they cannot link different identities when they are managing the database

Nykredit will be managing everything in-house, so it is difficult to prove that they cannot access the database and link the identities.

- It may be difficult for customers to adhere to the Nykredit technological standards

Nykredit may not be able to support all available technologies for their customers, so some customers, who are using a different setup than Nykredit, may find it difficult to comply with the Nykredit standard.

3.3.0.7 Database on 3rd Party Side

Advantages

- Neither companies nor Nykredit have to invest extra in IT infrastructure

The database is managed by the trusted 3rd party, who will invest in the infrastructure, so neither Nykredit nor the companies will have to invest extra in IT infrastructure

- It is easier for new and old customers to be a customer at Nykredit

The trusted 3rd party can support a wide range of technologies, so it is easier for customers to use their existing technology when becoming a new customer at Nykredit

- Nykredit can easily prove that it cannot link different Identities

Nykredit is not hosting the database, so it is easier for them to prove that they cannot link the identities.

- Data can still be retrieved in case the company stops existing

The database is always with the trusted 3rd party, so it does not matter if some company stops existing, the data can still be accessed. Special arrangements have to be made in case the trusted 3rd party ceases to exist, but this will be rare and in that case, Nykredit may decide to take over that part of the trusted 3rd party.

- Company cannot tamper with database

The companies do not have access to the database, so they cannot tamper with the data.

- Authorities only have to go to the trusted 3rd party to get the data in case its needed.

Disadvantages

- The 3rd party must be trusted by both Nykredit and its customers

The database is neither with the company nor Nykredit, so the external service provider should be trusted by both parties to hold their sensitive data.

- In case the trusted 3rd party goes out of business it might be difficult to retrieve the data.
- Companies do not have control over their own data

The database is maintained by an external service provider, so the companies have to store the data there and hence they do not have control over their own private data.

3.4 Current Banking System

The current banking system can be seen as following. There are 2 parts of the

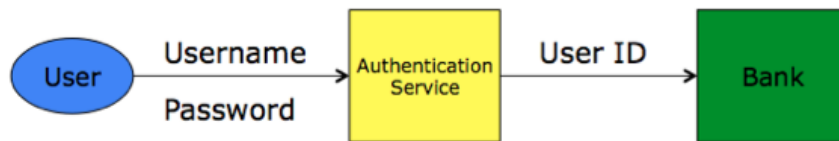


Figure 2: Current Banking System

Figure 3.2: Current Banking System

system:

- Authentication Service
- Bank

The end user interact with the service as follows:

- User goes to the authentication service and enters his credentials

- Authentication service authenticate the user and gives the bank User ID of the user
- All other details of the user are stored at bank side in relation with his user ID
- Bank provide services to the user

The authentication service can either be controlled by bank or a 3rd party (e.g. NemID)

In this system Bank is the most powerful entity. It has all the mappings

- User ID -> Account ID
- User ID -> Policy
- User ID -> User Personal Data
- User ID -> Transactions

This makes the bank most powerful identity in the system. Its very easy for bank to get any private data of customers using the data stored on bank side.

3.5 Summary

In nutshell current banking system gives bank a lot of power over user data. Also we have seen that traditional methods are not sufficient to provide proper anonymity to the users or they don't fit all the requirements properly.

CHAPTER 4

Proposed System

4.1 Introduction

This chapter will present our proposed system as the solution. We will define our system and show how it solves the problem of maintaining privacy for the users.

4.2 User Identification

In order to solve the problem we first look at how the users are identified in the system. There are 2 ways by which bank can identify the actual users

- From the logs i.e. transaction data

Banks store all the logs or transaction data with real id of the user. Its easy to access this data by the bank for a given user and extract all of his data.

- From a database in the bank system where personal details of the user are stored.

Bank stores personal data about all its users in a database. As this database lies at the bank, its possible for bank to use the database to get the personal information about a user.

We can remove this identification in following ways.

- Remove the user identity from the logs and transaction data

This way there is no way for the bank to get transaction data for a given user as there will be no user identity linked to the logs or transactions.

- Either remove the user personal data or limit the access to this personal database

If bank removes the personal database of the users from its side, then there is no way bank can get personal information of the users.

If bank limit access to such database and don't really use it for day to day banking purposes then its also possible for bank to limit access to user personal information.

4.3 Pseudonym System

In order to provide privacy to the users, we suggest a new pseudonym system. In

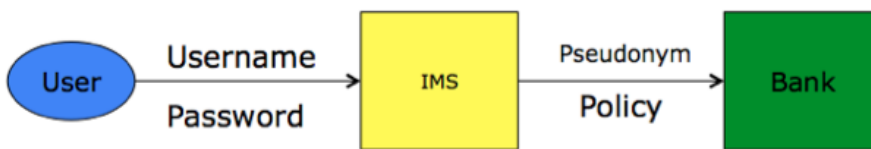


Figure 4.1: Pseudonym Banking System

our system we modify the Authentication service and replace it with an Identity Mapping System (IMS)

- Instead of giving the User ID to the bank we actually replace it with a pseudonym

- The IMS sends the policy as well as account ID to the bank
- Bank then use this information to provide service to the customer
- Bank doesn't need to store the mapping databases

The IMS can either be in the bank in a separate department or a 3rd party can manage it. This way the bank doesn't need to know the exact identity of the user to provide them with the services. Also bank can still store personal details of the user in case of legal requirement but it can be stored at a separate place as its not needed for day to day operation of the bank.

4.4 Properties

Following privacy properties are desirable in our pseudonym system.

- **Unlinkability** If the same pseudonym is used with different identities or different pseudonyms are used for same identity it should not be possible to link these different transactions to the same person.
- **Partial Information Disclosure** The information given by a user should be minimum and he should be able to choose what information values it actually wants to be made available to the bank.
- **Conditional Anonymity Removal** In case of some discrepancy or legal requirement, the authorities should be able to come in and identify the real user from the Pseudonym.
- **Revocation** It should be easy to revoke any user. Also it should be easy to check whether a certain user is revoked or not.

4.5 User Privacy

As in our system bank never gets the real identity of the user, the user anonymity to the bank is maintained. Also bank doesn't need to store the policies for the users as its all coming from the IMS. As a result, it decreases a lot of load from the bank to store such data. IMS service adds a layer of pseudonymity in the system.

4.6 Summary

In this chapter we presented our pseudonym system. Also we have given some certain privacy properties that our pseudonym system should be able to fulfill. In the next chapter, we will discuss about the prototype with the pseudonym system.

For our purposes we have decided to take case of 2 different systems for our IMS

- OpenID
- IDEMIX

CHAPTER 5

Simple Prototype

5.1 Introduction

In this chapter we will discuss about a simple prototype we made using the design from the pseudonym system in chapter 4. We will describe different parts of the system and how end user perceives it.

5.2 Design

The system is based on the design given in Chapter 4. For end user it doesn't

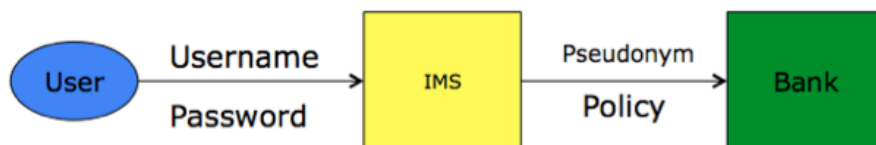


Figure 5.1: Prototype Design

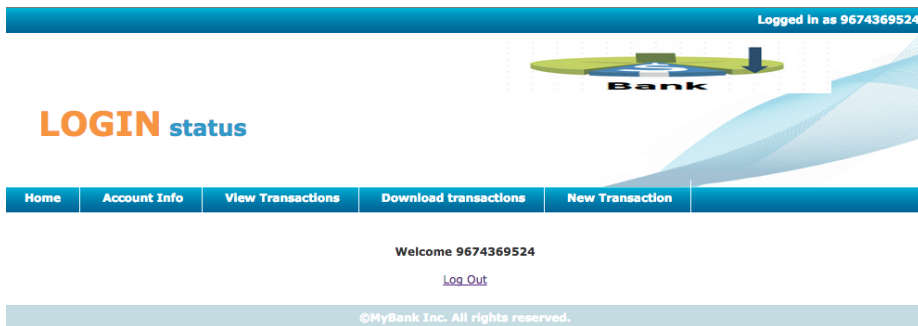
change anything. End user authenticates to the IMS and then IMS creates a pseudonym for the user. This pseudonym is then used by the bank for providing the services to the customer.

5.3 Authentication



The login page features a blue header with "Welcome to MyBank" on the right. Below the header is a navigation bar with links: Home, Account Info, View Transactions, Download transactions, and New Transaction. The main content area has a large "LOGIN" title in orange. To the right of the title is a graphic of a bank building on a green base with the word "Bank" below it. The login form includes fields for "Your Account ID:" (with "123" entered), "Your username:" (with "shanky" entered), and "Your password:" (with "*****" entered). A "Login" button is at the bottom of the form. The footer contains the text "©MyBank Inc. All rights reserved."

Figure 5.2: User Login Page



The authenticated user page features a blue header with "Logged in as 9674369524" on the right. Below the header is a navigation bar with links: Home, Account Info, View Transactions, Download transactions, and New Transaction. The main content area has a large "LOGIN status" title in orange. To the right of the title is a graphic of a bank building on a green base with the word "Bank" below it. The page displays "Welcome 9674369524" and a "Log Out" link. The footer contains the text "©MyBank Inc. All rights reserved."

Figure 5.3: Authenticated User


User authentication happens as follows:

- User goes to the bank login page.
- User puts his credentials in the login system.
- User credentials are verified by IMS and then user is redirected to his account page.

One thing to note that is that as traditional user, this doesn't change anything on the user end. User still use the same process to get access to his account.

As we can see in 5.3, after authentication user is logged in with a pseudonym.

5.4 User Information



Logged in as 2322497440

Account INFO

Home Account Info View Transactions Download transactions New Transaction

Mybank have following information about you

Name	Value
Pseudonym	2322497440
Account ID	123
Balance	6870
Account Type	Corporate
Withdraw Permission	1
Transfer Permission	1

©MyBank Inc. All rights reserved.

Figure 5.4: User Information - Session 1

When we go to user information page we can see the information that is given to the bank by IMS for a given user.

In our case it is:

- Pseudonym
- Account ID
- Balance
- Account Type
- Policies
 - Withdraw Permission



Figure 5.5: User Information - Session 2

– Transfer Permission

As we can see from 5.4 and 5.5, 2 different sessions of the same user are logged in with different pseudonym. Bank have no way to find out that its the same user who have logged into different sessions.

5.5 Transactions


In our prototype user is allowed to do 2 types of transactions

- Debit
- Credit

All the transactions that are done by the user are logged in with the pseudonym; with which user has been logged in to the system.

5.6 shows the new transactions page in the system where user is allowed to do the transactions.

Logged in as 5406571313



New Transaction

Home

Account Info

View Transactions

Download transactions

New Transaction

Amount: * 100

Type: * Debit

Details: * Screenshot

*Required

Submit

©MyBank Inc. All rights reserved.

Figure 5.6: New Transactions

5.7 shows all the transactions that has been done on the given account by users. As we can see, all the transactions are saved with the pseudonym of the users.

Our system also allow the users to download the transactions from the download transactions page as shown in 5.8. These transactions are stored in a csv file and then can be seen as in 5.9.

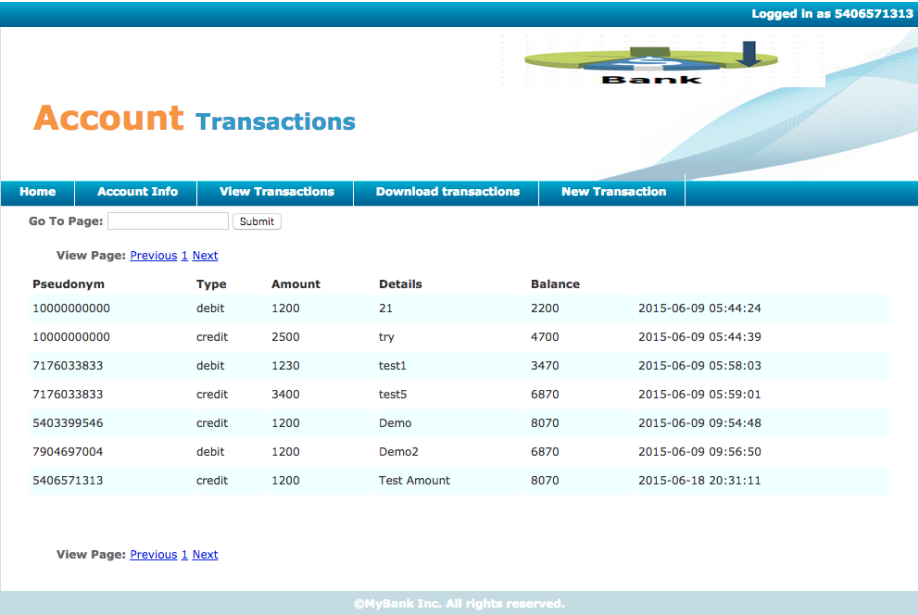


Figure 5.7: Account Transactions

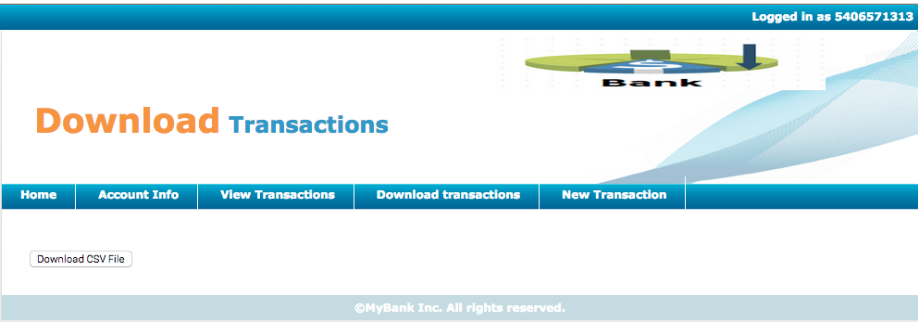
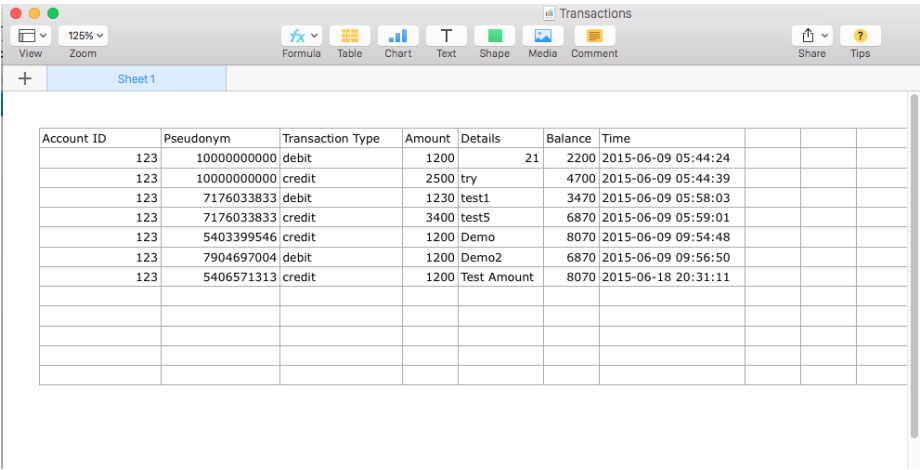


Figure 5.8: Download Transactions Option



The screenshot shows a spreadsheet application window titled "Transactions". The interface includes a menu bar with options like View, Zoom (125%), Formula, Table, Chart, Text, Shape, Media, Comment, Share, and Tips. Below the menu bar is a tab labeled "Sheet 1". The main area contains a table with the following data:

Account ID	Pseudonym	Transaction Type	Amount	Details	Balance	Time			
123	10000000000	debit	1200	21	2200	2015-06-09 05:44:24			
123	10000000000	credit	2500	try	4700	2015-06-09 05:44:39			
123	7176033833	debit	1230	test1	3470	2015-06-09 05:58:03			
123	7176033833	credit	3400	test5	6870	2015-06-09 05:59:01			
123	5403399546	credit	1200	Demo	8070	2015-06-09 09:54:48			
123	7904697004	debit	1200	Demo2	6870	2015-06-09 09:56:50			
123	5406571313	credit	1200	Test Amount	8070	2015-06-18 20:31:11			

Figure 5.9: Downloaded Transactions File

5.6 Summary

In this chapter we showed a simple prototype of our system. This system works with pseudonyms and in the next chapter we will talk about how we can replace our Identity Mapping System with OpenID and IDEMIX for our purposes.

CHAPTER 6

OpenID Based Solution

6.1 Introduction

In this chapter we will provide a solution using OpenID based IMS. We will give more details about how this system can be implemented, and how will it behave for the end users.

6.2 OpenID IMS

We can replace IMS with OpenID based IMS in our pseudonymous system. This system will take user credentials and then send Pseudonym, Account ID and Policy to the bank. This IMS can be controlled by a separate identity inside the bank or by a 3rd party.

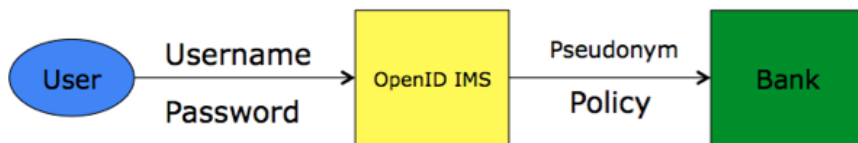


Figure 6.1: Pseudonym System with OpenID IMS

6.3 System Setup

The system can be setup in 2 ways:

- IMS controlled by a separate department at bank.
In this case the bank separate the authentication and service part in 2 different departments internally. Authentication is controlled by IMS which holds the sensitive user data but the service department doesn't need to have access to that data to provide service to the user.
- IMS controlled by a third party
In this case the bank operates the service part while the authentication part is operated by a trusted third party.

In both cases, bank and IMS have to collaborate and bank have to trust the IMS system that pseudonym and policy sent by the IMS system is correct.

6.3.1 Changes on the Bank Side

In order to provide services to a pseudonym bank needs to have a temporary policy database at its end. So that when bank gets the pseudonym and policy from the IMS system, it can store that policy with the pseudonym and provide the services according to the policy.

6.3.2 Information Stored at IMS

IMS need following user information to be stored

- User ID

- Account ID
- Policy

Account ID and policy can be stored in encrypted form, which can then only be opened by the bank. OpenID IMS also need to store a mapping database from User ID to Pseudonym for escrow purposes.

6.3.3 Changes needed on the User Side

On user side no changes are needed. User access the system like before. User doesn't need to install any special software or hardware on his side to access the bank services.

6.4 User Creation

Following are the steps for creation of a new user account in OpenID based system

- User goes to bank to open a new account.
- User provides his details.
- Bank creates user policy and send this information to the IMS system along with other user information.
- IMS system verifies the user information and provides user with credentials to access his account.
- User then can login to his account using the credentials.
- In case of corporate users, if user is the administrator then he can add more users using a web interface at the IMS system directly and decide account policies for those users.

6.5 User Authentication

Authentication steps are as following in OpenID based system

- User goes to login page
- User provides his username and password
- This is sent to OpenID IMS which verifies the user and creates a pseudonym for the given user ID
- This user ID to pseudonym mapping is stored in the database for escrow purposes
- IMS gets the policy for the given user ID from the policy database
- IMS then sends the pseudonym, Account ID and Policy to the bank
- Bank gets this information and create a temporary policy for the given pseudonym
- User can then access services from the bank using the pseudonym
- All user transactions are logged with the pseudonym

6.6 ID Escrow

Following are the steps for ID escrow in OpenID based system

- Authorities come to the bank for transaction data.
- After verifying, bank gives the transaction data to the authorities.
- Authorities then go to the IMS based system for the mapping data.
- After verifying, IMS gives the mapping data to the authorities.
- Authorities then get the real identity of the user from mapping and transaction data.

6.7 Summary

This chapter described the IMS system setup using the OpenID system. We described how the system will be setup and how it will affect all the parties involved. We will analyze this system later in chapter 8.

CHAPTER 7

IDEMIX Based Solution

7.1 Introduction

In this chapter we will provide a solution using IDEMIX based IMS. We will give more details about how this system can be implemented, and how will it behave for the end users.

7.2 IDEMIX IMS

As in previous case we can replace IMS with IDEMIX based IMS in our pseudonymous system. This system will take user credentials and then send an IDEMIX token to the bank. This IDEMIX token contains pseudonym as well as account ID and policy for the user. This IMS can be controlled by a separate identity inside the bank or by a 3rd party.

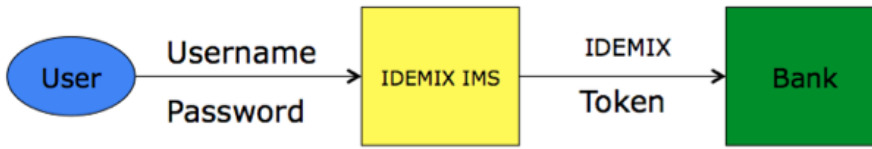


Figure 7.1: Pseudonym System with IDEMIX IMS

7.3 System Setup

The system can be setup in 2 ways:

- IMS controlled by a separate department at bank.
In this case the bank separate the authentication and service part in 2 different departments internally. Authentication is controlled by IMS which holds the IDEMIX credentials for the user. Service department only gets the IDEMIX token from the authentication department.
- IMS controlled by a third party
In this case the bank operates the service part while the authentication part is operated by a trusted third party.

In both cases, bank and IMS have to collaborate. Bank have to trust the IMS system that IDEMIX token sent by the the IMS system is correct.

7.3.1 Changes on the Bank Side

In this system, bank needs to behave as an IDEMIX issuer and verifier. It will issue IDEMIX credentials for the users and also will verify the tokens sent by the IMS system.

7.3.2 Information Stored at IMS

IMS system will behave like user in the IDEMIX system. IMS needs following user information to be stored

- User IDEMIX credential

Account ID and policy can be stored in encrypted form in the credential itself. This IDEMIX credential for a particular user can be setup in the beginning and then can be used later to create authentication tokens.

7.3.3 Changes needed on the User Side

On user side no changes are needed. User access the system like before. User doesn't need to install any special software or hardware on his side to access the bank services.

7.4 User Creation

Following are the steps for creation of a new user account in OpenID based system

- User goes to bank to open a new account.
- User provides his details.
- Bank creates user policy and send this information to the IMS system along with other user information as an IDEMIX credential.
- IMS system verifies the IDEMIX credential for the user and provides user with credentials to access his account.
- User then can login to his account using the credentials.
- In case of corporate users, if user is the administrator then he can add more users using a web interface at the bank directly and decide account policies for those users.

7.5 User Authentication

Authentication steps are as following in IDEMIX based system

- User goes to login page
- User provides his username and password

- This is sent to IDEMIX IMS which then gets the saved user credential and creates a presentation token with a pseudonym for the bank
 - Also for escrow purposes the real user identity is also encrypted in the token with the public keys of authorities
- Bank receives this presentation token and gets the following information
 - Pseudonym
 - Account ID
 - Policy
- Bank adds this information in a temporary policy database
- Bank saves the token for future escrow purposes
- User can then access services from the bank using the pseudonym
- All user transactions are logged with the pseudonym

7.6 ID Escrow

Following are the steps for ID escrow in IDEMIX based system

- Authorities come to the bank for transaction data and IDEMIX token.
- After verifying, bank gives the transaction data and corresponding IDEMIX token to the authorities.
- Authorities then using their key get the real identity of the user from the IDEMIX token.

7.7 Summary

This chapter described the IMS system setup using the IDEMIX system. We described how the system will be setup and how it will affect all the parties involved. We will analyze this system later in chapter 8.

Analysis

8.1 Introduction

This chapter will present the analysis of the 2 systems presented in Chapter 6 and 7. Also we will take the better system according to our needs and present the modified system according to our system.

8.2 OpenID Based pseudonym System

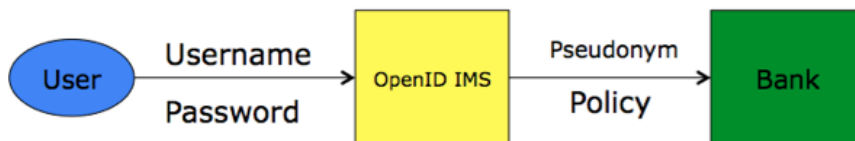


Figure 8.1: Pseudonym System with OpenID IMS

With the use of OpenID IMS we add a pseudonymous layer in the system. This

provides us the necessary privacy. But in order to do so OpenID provider needs access to a lot of data. Some of the example data is:

- UserID
- Account ID
- Policies

In addition to that, the provider needs to store the mapping database from UserID to Pseudonym. Bank really has to trust the provider with storage of all this sensitive data. In some cases bank might not want the provider to store such data by themselves.

In case there is a discrepancy, the authorities need to go both to the bank to get the transaction data as well as the provider for the mapping data.

8.3 IDEMIX Based pseudonym System

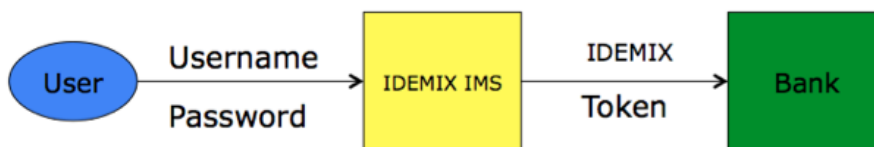


Figure 8.2: Pseudonym System with IDEMIX IMS

With the use of IDEMIX IMS we add a pseudonymous layer in the system. This provides us the necessary privacy. In order to do so, IDEMIX IMS just need to store the IDEMIX credential of the user.

The provider doesn't need to store any mapping database on his side. It is easier for bank to implement, as bank really doesn't have to trust the IDEMIX IMS to store sensitive data.

In case there is a discrepancy, the authorities need to go only to the bank to get the transaction data as well as the mapping data from the IDEMIX tokens.

8.4 IDEMIX implementation in the Real World

From above 2 analyses, we conclude that IDEMIX implementation of pseudonymous system is more favorable to bank than the OpenID implementation. Now we will try to fit this implementation in our system, which includes Nykredit as the Bank, Signicat as the 3rd party, DTU as corporate customer and other government institutions as authorities.

8.4.1 Addition of the New User

Addition of the new user can happen as following:

1. DTU registers the new user with the Nykredit giving them the user details and policies that should apply to the particular user regarding the account.
 - (a) Nykredit registers this new user with his User ID with the IMS system
2. Nykredit issue an IDEMIX policy credential for the given user to DTU. This credential contains the policy information and account information for the user.
3. DTU then use this policy credential to register the new user with Signicat.
 - (a) Signicat inquire about the user data with the authorities
 - (b) Authorities verify the user data to Signicat
4. After that Signicat issue final IDEMIX credential for the IMS system. This credential is then used to create pseudonym IDEMIX tokens for the user.

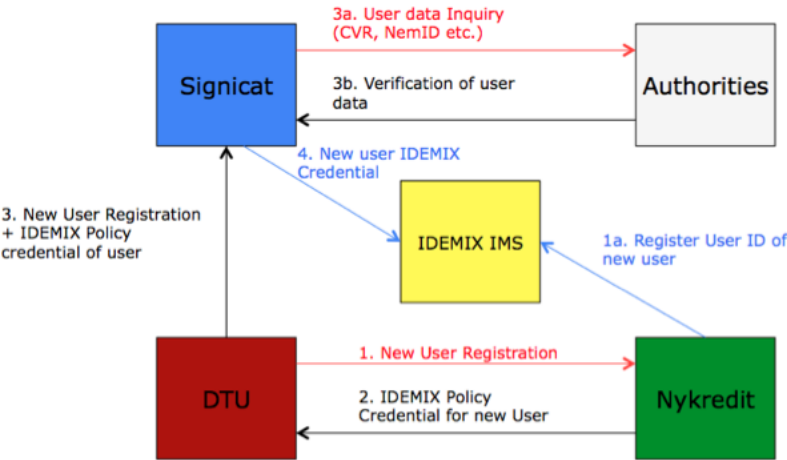


Figure 8.3: IDEMIX Credential issuance for a new user

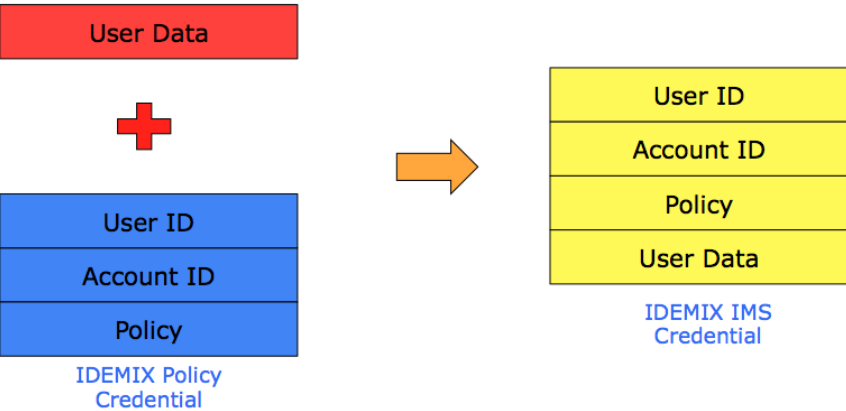


Figure 8.4: Final IDEMIX Credential from Policy Credential

8.4.2 Addition of a New Customer

Addition of the new customer is almost same as addition of new user:

1. An administrator goes to Nykredit to open a bank account on behalf of DTU
 - (a) Nykredit registers DTU as new customer in their internal system.
 - (b) Nykredit register the DTU administrator with his User ID with the IMS system
2. Nykredit issue an IDEMIX policy credential for the DTU administrator to himself. This credential contains the policy information and account information for the administrator.
3. Administrator then use this policy credential to register himself as owner of the new DTU account with Signicat.
 - (a) Signicat inquire about the data given in the credential with the authorities
 - (b) Authorities verify the data to Signicat
4. After that Signicat issue final IDEMIX credential for the IMS system. This credential is then used to create pseudonym IDEMIX tokens for the administrator.

8.4.3 Technical Requirements

In this system DTU as a client doesn't need to change anything on their side to be a customer at Nykredit. All the system for DTU is web based where they can just add/remove users and also DTU users login to the system using the browser.

Nykredit have to implement IDEMIX issuer service on their side to issue IDEMIX Policy credential. This is done so that Nykredit doesn't have to store the sensitive data at the 3rd party. Use of this credential ensures that this data remains safe. Nykredit also have to implement IDEMIX verifier service to verify the user identity.

Signicat have to implement IDEMIX issuer service also to issue final IDEMIX credentials.

IMS have to implement IDEMIX user service to create the IDEMIX tokens for the user while user is logging in.

8.5 Summary

In this chapter we analyzed 2 different pseudonym systems as discusses in chapter 6 and 7. Also after that we discussed in detail the implementation of the IDEMIX system in the real world in our system.

CHAPTER 9

Conclusion and Future work

9.1 Introduction

This chapter will conclude the thesis and will provide directions for the future work that can be carried out in the field.

APPENDIX A

Appendix A

This appendix is full of stuff ...

Bibliography
