Data protection in the Cloud – The MimoSecco approach

Jonas Lehner, Andreas Oberweis, Gunther Schiefer

Karlsruhe Institute of Technology
Institute AIFB

Jonas.Lehner@kit.edu, Andreas.Oberweis@kit.edu,
Gunther.Schiefer@kit.edu

Keywords: Cloud Computing, Data protection, Context-based access control, Legal Aspects, Software as a service, Secure hardware

Abstract. Cloud Computing is a technology with vast impact on IT systems. Costs can be significantly reduced through on-demand purchase of CPU time, memory and storage, offering high flexibility. The main reason to avoid cloud technology still is security. This leads to a lack of trust in cloud services. Most cloud providers secure their systems only against external adversaries by using firewalls and secure connections. Internal adversaries, however, remain a big threat in this scenario. Especially when using mobile devices as clients, usable security with a low performance impact remains a challenge.
In this paper, we present concepts for using software as a service with mobile devices while guaranteeing a high level of data protection. MimoSecco uses an innovative encryption scheme and hard-to-clone secure hardware to guarantee data protection. Top secret data is encrypted directly, processible confidential data is encrypted and fragmented by the database proxy and transferred to different servers. Context-based access control makes the misuse of mobile devices for unauthorized data access difficult. These set of measures raises the privacy level of cloud computing significantly.

1 Introduction

1.1 Motivation

For many small and medium companies, cloud computing is very attractive. It offers scalability to cope with temporal high server loads. Building and maintaining an own computing center is very costly and can be avoided by using flexible cloud services. Therefore, costs can be cut while outsourcing maintenance to specialized cloud providers. Overall, this leads to an improved use of the available resources and lower costs. Especially in the last years, more and more mobile devices like smartphones or tablets are bought. While offering high mobility, they often lack in computing power. Cloud providers like Google respond by moving complex calculations or big data

adfa, p. 1, 2011.

bases into the cloud, while using the mobile device only for the user interface and other simple tasks. This allows the devices to run longer while still offering the same service.

Once IT systems are outsourced into the cloud, the user loses control over his data. Using cloud services always leaks sensitive information, resulting in huge privacy issues but offers the possibility to access data from anywhere without the necessity of operating costly infrastructure for this purpose.

Protection against external adversaries is not enough. Especially so-called insider attacks have to be taken into account when designing a cloud service that uses a public cloud. Usually administrators have lots of privileges that often allow them to read the client's data, e.g. for making backups.

Data privacy laws, however, are impossible to hold for the big US cloud providers (Google, Amazon, Microsoft …), but also Chinese and Russian providers. In US, for example, authorized through the patriot act, the government can access all data of companies located in the US, even if stored outside the US. This also holds for data centers within the European Union.

1.2 Overview

MimoSecco aims at solving the cloud privacy dilemma by encrypting outgoing sensitive data before it is stored in the potentially untrusted cloud. We present a software-based solution with the option to integrate secure hardware for further security improvements. This article focuses on data protection in the terms of confidentiality, this means to avoid any information retrieval by a person which is not authorized to do so.

Other security problems like data loss, availability, integrity and so on are not discussed here; most of them can be solved by system and/or data redundancy. A detailed description of related work and the technical architecture of the systems can be found in [[VERWEIS AUF PAPER 36]]. Side-channel attacks to MimoSecco are discussed in [[VERWEIS AUF PAPER VON Matthias Huber]].

For a better understanding it is proposed to read [[VERWEIS AUF PROJEKTÜBERBLICK MimoSecco]] and [[VERWEIS AUF PAPER 36]] first. This article focusses on additional mechanisms and legal aspects of MimoSecco.

1.3 Project Goals

The aim of the project MimoSecco is to protect data, processed by a cloud provider, better against internal attacks. Internal adversaries are all persons who have legal access to the rooms or the IT systems of a cloud provider. This covers, for example, the administrator, the service technician and the cleaning personal. Protection against internal adversary includes automatically the protection against external adversaries. To do this, MimoSecco uses encryption technologies and separates the storage of the data from the processing. This reduces, for example, the risk, that data could be stolen by taking away a backup tape, because the provider who processes the data does not have a backup and the provider who stores the data cannot decrypt it.

By the integration of context based access control it is possible to grant data access only when specific context parameters are given. One example is that you are not allowed to access confidential technical documents while your calendar shows that you are on holiday.

2 Architecture

The MimoSecco architecture is derived from the outsourcing setting using cloud services (cf. Fig. 1). The hardware of the data owner is assumed completely trustworthy. The storage cloud is considered an honest-but-curious (passive) adversary. This means all queries are answered correctly, but the cloud provider tries to gain information from the queries and stored data. Between the trusted zone and the cloud storage provider is the semi-trusted zone. It offers better privacy due to a more favorable jurisdiction than US-based providers or is provided by a known and trusted business partner. A private cloud is one example for this zone.

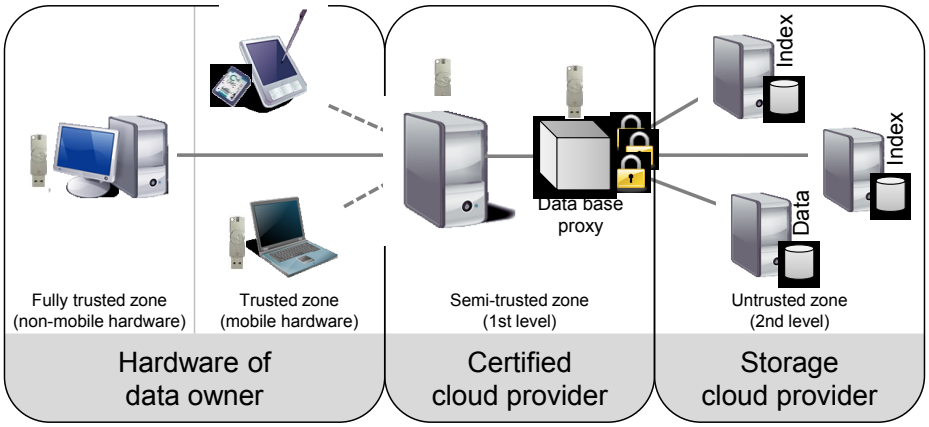

Fig. 1. Security zones in the MimoSecco scheme.

Most cloud security solutions only focus on transport encryption. MimoSecco uses authenticated end-to-end-encryption, too. Not only is the server authenticated, like in common cloud-based services, but also the client. The certificates for this are generated within the company (fully trusted zone) and stored securely on a state-of-the-art smart card. We use the *CodeMeter*[1] platform of WIBU-SYSTEMS AG for secure storage. CodeMeter is a USB/SD/… dongle for software protection and offers password-protected storage.

MimoSecco also takes the potentially curious cloud provider into account and does not only rely on encrypted connections. Cloud providers pose high-value targets since they host services for multiple companies. If an adversary manages to corrupt such a

[1] For more information about CodeMeter visit http://www.wibu.com/en/codemeter.html

provider, he can access the data of a potentially large number of clients. MimoSecco prevents damage in such a case by encrypting the data before uploading it to the second level cloud provider. The decryption is only possible with authentication by the user. The cryptographic keys required for encrypting and decrypting of data are distributed via the WIBU-SYSTEMS CodeMeter License Central[2] between the data owner's smart card and the cloud provider's smart card. MimoSecco only needs CodeMeter in the user's zone and for the certified cloud provider. Since a considerable contract between controller (user) and processor (certified cloud provider) is needed, the usage of a smartcard on the processor's server is a minor problem. Additionally, the ability to offer such an increase of security is a competitive advantage for the processor.

If a requested operation on the stored data is authenticated by the data owner and the Software as a service application (necessary for access control, see chapter 4) the smart card of the database proxy enables the proxy to use the cryptographic keys to encrypt or decrypt the data as needed.

3 Security Levels

One goal of MimoSecco is that a user retains control over his data. To ensure this, data has to be classified in different security levels, which is derived from *mandatory access control* (MAC, see chapter 4). Fig. 2 shows the different data channels for the security levels.

Top secret data (black) is encrypted by the user's smart card on his device. This data can't be decrypted by the cloud provider. Therefor top secret data can't be processed by the cloud provider but can only be stored. The secret key for this type of data stays strictly at the user.

Public data (gray) is not encrypted at all. This type of data is not worthy of protection or accessible easily from other source, e.g. published documents from the internet. This kind of data is not considered since it is not interesting in the MimoSecco context.

[2] More information about CodeMeter License Central can be found at http://www.wibu.com/en/online-software-activation.html

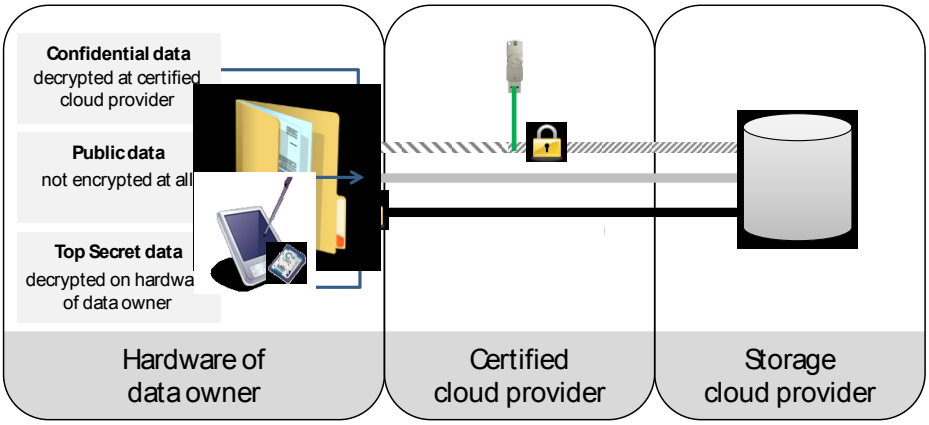

Fig. 2. Security levels of MimoSecco architecture

Confidential data (striped) is the mainly regarded kind of data in MimoSecco. To profit from the advantage of cloud computing these data has to be processible by the cloud provider who therefor needs to be able to decrypt the data, at least for a short period of time. The mechanism for decryption and encryption at this point and the technical implications are described in [[VERWEIS AUF PAPER 36]].

4 Context-based Access Control

A characteristic feature of the MimoSecco approach is the context-based access control. The task of an access control in a computer system is to decide whether a user's request to execute a certain operation on a certain resource can be accepted [1]. Examples for resources are files, database objects or services (web services, print services etc.). Operations are commonly *read*, *write* and *execute*. The user is the active element and is also called subject, while the resource as the passive element is called object. *Permission* combines an object with an operation.

Access control models can be separated into generic models and application specific models. The latter are developed to be used in a certain application context like in a database management system or in a workflow management system whereas generic models don't have a defined context. There are three different relevant generic models for access control described in literature: mandatory access control (MAC), discretionary access control (DAC), and role-based access control (RBAC).

Mandatory Access Control (MAC).

This model is mainly used in high security areas like military or intelligence and is based on the principle of different trust levels (e.g. public, confidential, secret or top secret) [2]. These trust levels are assigned to both users and resources. To access a certain resource a user has to have at least the same trust level (e.g. to access a *secret* resource, a user needs to be either *secret* or *top secret*).

Discretionary Access Control (DAC).

The principle of this access model is an access matrix, which contains the rights of each user to access each resource [3; 4]. A user, who creates an object (e.g. data) is initially entitled with all operation rights (e.g read, write or execute). He can then grant other users some or all of these rights. Users can be grouped to simplify configuration.

Role-based Access Control (RBAC).

RBAC uses roles to define access rights [5]. Roles correspond with task descriptions for jobs or positions in the organization (e.g. *marketing manager* or *board member*). Access rights are never assigned to single users but always to roles.

In MimoSecco we took a DAC model as a basis since it is easy to implement yet still powerful enough to satisfy the project's needs. We then added a context aspect.

Context in MimoSecco is understood as dynamic information that is explicitly available during the runtime of the system and that can be used to adjust the application to the user's situation [6; 7].

Context-sensitivity means that permissions depend on context information, e.g. location, time or calendar data. For example, if a sales representative is located on the company area of a certain customer, he is not allowed to access data of a competitor.

To implement the context-sensitivity, we use so called context switches, which get context information as an input and make decisions based on rules defined by the company. As shown in Fig. 3 the context switches can affect the access control decision at two different points:

- **Stand-alone permission**: Permissions can be switched on and off according to actual context situation (e.g. reading of certain data is forbidden when using an unsecure wireless LAN connection)
- **Type of object**: In certain situations particularly sensitive data can be protected against some operations. I.e. no specific record is protected but an entire type of object (e.g. an employee can't access personal files while he is not within the company area).

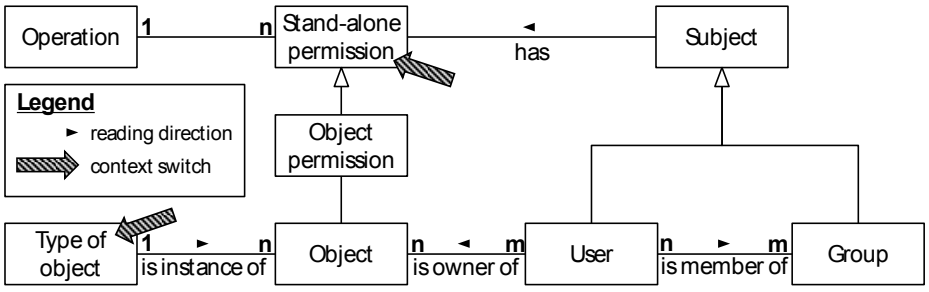

Fig. 3. Context-based discretionary access control model used by MimoSecco

Some of the relevant context parameters we consider in MimoSecco are:

- **Location**: To protect sensitive data against access at places, which don't offer the required privacy, e.g. public airports, the mobile device can be located via GPS or interpretation of the IP address.
- **Time**: The context parameter time can be used to prevent access to sensitive data out of office hours. For global companies it can be necessary to determine the local time of an employee by using location data. In this case one context parameter is used as an input value for another context parameter.
- **Calendar data**: Calendar records can be used to determine whether an employee should have the permission to access certain data. For example, during his holidays an employee is not allowed to access any data.
- **Type of authentication**: Depending on the type of authentication a user can get different permissions. It is taken into account whether a user uses a username and password or if he is using a smartcard or hardware token.
- **Type of connection**: Different connection types can carry different risks that are regarded for the decision.

5 Points of attack

In MimoSecco we focus on so-called insider attacks. If we protect the data against adversaries inside the companies, this is also effective against adversaries from outside, therefore the arrows in the next Figure are striped gray-black. As shown in Fig. 4 there are four points of attack in our model.

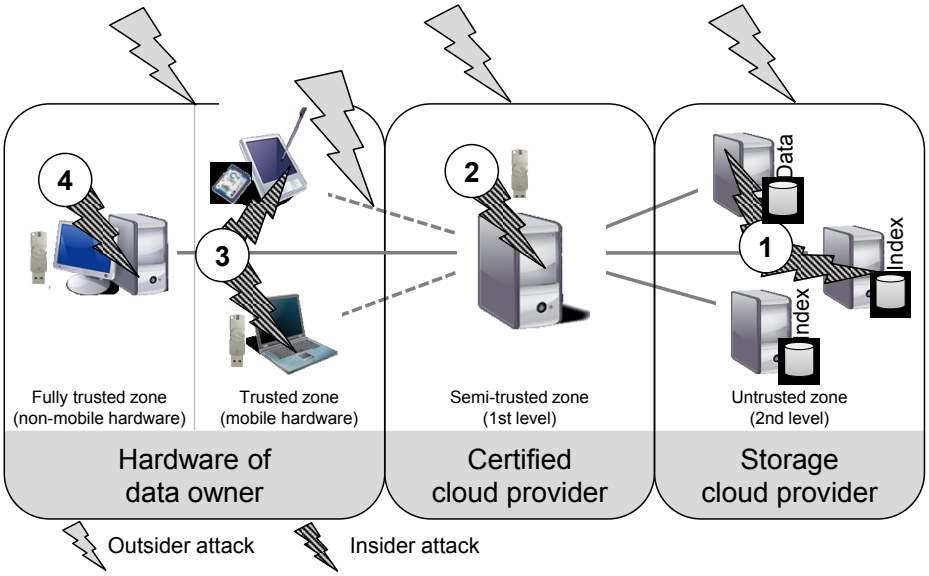

Fig. 4. Points of attack in the MimoSecco scheme

(1) The 2nd level cloud provider is the host for all data. He only gets encrypted data (except the indices, where the keywords are in plain text). Therefore, nobody there can gain Information about the data. Any adversary can only learn something about the keywords which must be somewhere in the whole database, but he gets no information about the relationships.

(2) Nowadays, the certified cloud provider in the semi-trusted zone needs the decrypted data, to deal with it. (This can change in future). Therefore it is necessary to temporary decrypt the data in the memory. Since the data is not stored unencrypted in this zone and the keys are also only in memory during use, it is more difficult (but not impossible) to get some information here. It is feasible to implement some methods to control the data-usage, which avoids the reading of the whole database. The TrustedCloud-project "Sealed Cloud" offers a solution to get more protection against an adversary at this point. For further information about Sealed Cloud see [[VERWEIS AUF PROJEKTÜBERBLICK Sealed Cloud]]

(3) The context-based access control and the usage of a hardware security token makes it significantly more difficult, to misuse the data on a mobile computer. There is no data stored on a mobile device, except caching-data, which is automatically deleted after use.

(4) The possibility to misuse the data by a person, who has the right to deal with it, is not changed through the MimoSecco model. Since this problem is independent from the usage of cloud computing it is not considered here.

By using cloud computing, the data is not persistent stored on the computers of the cloud user. Also the certified cloud provider does not store the data. Since there is no persistent data, there is nearly nothing to read, delete or modify for an adversary. The usage of a sufficient transport encryption is a matter of course and independent of wired or mobile connections.

6 Legal Aspects

One goal of MimoSecco is to enhance data security. Additionally there are a lot of use cases, where one uses personal data, for example in a CRM scenario where an enterprise stores information about their customers using a cloud service. This makes it necessary to deal with the questions concerning to personal data.

The upcoming *General Data Protection Regulation (GDPR)* from the European Union is still a draft. It can take several years since it becomes effective. Until then, the German *Bundesdatenschutzgesetz (BDSG)*, which is based on the European Data Protection Directive 95/46/EC, is the main basis for dealing with personal data (in Germany).

In the MimoSecco model, there is a certified provider that offers a cloud service as SaaS (semi-trusted zone). This provider is a processor (Auftragnehmer) in the terms of the BDSG. According to the BDSG it is necessary to set out a contract in writing between the controller (Auftraggeber) and the processor. If the controller and the processor follow the full terms of § 11 BDSG, the processor is allowed to act with the

personal data as if he is part of the controller. This proceeding is only allowed between an enterprise and a provider within the European Union (or some especially allowed countries).

For data storage, MimoSecco uses the database proxy to encrypt the data and store it within further cloud providers. These storage providers are set in the untrusted zone. This means, that the certified provider transfers the data to another party, the storage provider. According to the BDSG, the transfer of personal data is not allowed in most cases. In MimoSecco we have the transfer of highly encrypted data. The open question is, whether encrypted data is still personal data or not, since the transfer of personal data is mostly not allowed. To detail this question, it has to be determined, if encrypted data is pseudonymous or anonymous data. The aim of making data pseudonymous is to obscure the identity of personal data, but to have the possibility to re-identify the concerned person. The goal of anonymisation, however, is to make it impossible for everyone to re-identify the concerned person. According to this, anonymised data is no longer personal data, since pseudonymised data is still personal data. The Goal of anonymisation described here doesn't match with the definition which is given in § 3 (6) BDSG. The BDSG defines anonymisation as modifying of personal data in a way, that it is not possible to re-identify a concerned person or that it needs a disproportional effort of time, costs and manpower to do that. This last part of the definition matches with the result of a strong encryption, which is a strict version of pseudonymisation.

To answer the given question, the point of view is relevant. This has to be divided into a bifocal perspective (e.g. Stiemerling and Hartung [8]). From the subjective view of a provider, who receives encrypted data, this is no personal data at all, since he is not in possession of the decryption key (or any other possibility) to decrypt the data. It doesn't matter, if the original (unencrypted) data was personal data or not. From an objective view it is not possible to convert personal data through encryption into non-personal data. This is because there is still someone who has a key and the possibility to decrypt the data and to restore the personal information. Another open question is, if encrypted personal data is no longer personal data, when all keys to decrypt it are destroyed. Since this is a separate question, it isn't discussed further here.

According to Stiemerling and Hartung is the mostly in Germany followed opinion, the subjective view. This means that encrypted personal data has not been treated as personal data by the provider, if he doesn't have the key at his disposal. There are still no court decisions which clarify this.

The GDPR (also directive 95/46/EC) follows the objective view. The given reason (23) at the beginning of the GDPR says: "The principles of protection should apply to any information concerning an identified or identifiable person. To determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or *by any other person* to identify the individual. The principles of data protection should not apply to data rendered anonymous in such a way that the data subject is no longer identifiable." From this point of view it is not possible to convert personal data in non-personal data in any way (as long as a key to decrypt it exists). From this point of view, a 1st level provider (semi-trusted zone)

should use only storage providers (2nd level provider) within the European Union (or provider which have the same legal status according to the GDPR) and create in each case a relationship - between the 1st level provider as a controller and a storage provider as processor - according to the full terms of § 11 BDSG. This restricts the world wide cloud to a European cloud.

As a conclusion it could be said, that (following the dominant opinion in Germany) it is legal for a German provider to use the MimoSecco model for storing data on 3rd party clouds without following the rules of § 11 BDSG between the semi-trusted and untrusted cloud provider. Since there are no court decisions about this, there is no legal certainty. For a careful, privacy-aware enterprise, which wants predictability of legal decisions, the MimoSecco model should only be used in a European cloud as described above until further clarification of the legal situation.

7 Conclusion and Future Work

In this paper, we introduced a method to solve the cloud dilemma. With the MimoSecco transformation, outsourcing of sensitive information to the cloud is possible.

Still, numerous challenges lie ahead. The legal aspects have to be clarified to give companies legal compliance. Other security problems like data loss, availability, integrity and so on have to be regarded. Some of them can be solved by system and/or data redundancy. This needs further research, which should also consider the performance of the whole system.

Another open question is the long time security of encrypted data. Nobody knows how long data, encrypted according to the state of the technology, is secured. How about a backup made by a storage cloud provider, which rests in a shelf for decades? Another planned feature is the use of the new German ID card (*neuer Personalausweis, nPA*) for authentication.

This work has been partially funded by the Federal Ministry of Economics and Technology, Germany (BMWi, Contract No. 01MS10002). The responsibility for the content of this article lies solely with the authors.

1. S.D.C. Vimercati, S. Paraboschi, P. Samarati: Access Control: Principles and Solutions. Software — Practice and Experience, vol. 33, no. 5, 2003, p. 397-421.
2. M. Benantar: Mandatory-Access-Control Model. Access Control Systems: Security, Identity Management and Trust Models, 2006, p. 129-146.
3. B.W. Lampson: Protection. Operation Systems Review, vol. 1., no. 8, 1974, p. 18-24.
4. R.S. Sandhu, P. Samarati: Access control: principle and practice, Communications Magazine, IEEE , vol.32, no.9, Sept. 1994, p. 40-48.
5. R.S. Sandhu, E.J. Coyne, H.L. Feinstein, and C.E. Youman: Role-based access control models. Computer, 29(2), 1996, p. 38-47.
6. A.K. Dey: Understanding and Using Context. Personal and Ubiquitous Computing Journal, vol. 5, no. 1, 2001, p. 3-7.
7. G. Chen, D. Kotz: A Survey of Context-Aware Mobile Computing Research. Technical Report TR2000-381, Department of Computer Science, Dartmouth College, Hanover, NH, USA, 2000.

8. O. Stiemerling, J. Hartung: Computer und Recht: Zeitschrift für die Praxis des Rechts der Informationstechnologien, Köln, 1/2012, p. 60-68.