Project N°: **FP7-284731**

Project Acronym: **UaESMC**

Project Title: **Usable and Efficient Secure Multiparty Computation**

Instrument: **Specific Targeted Research Project**

Scheme: **Information & Communication Technologies**

Future and Emerging Technologies (FET-Open)

Deliverable D1.2
Requirements specification based on the interviews

Due date of deliverable: 31st July 2012

Actual submission date: 31st July 2012

Start date of the project: **1st February 2012** Duration: **36 months**

Organisation name of lead contractor for this deliverable: **UT**

Specific Targeted Research Project supported by the 7th Framework Programme of the EC		
Dissemination level		
PU	Public	✓
PP	Restricted to other programme participants (including Commission Services)	
RE	Restricted to a group specified by the consortium (including Commission Services)	
CO	Confidential, only for members of the consortium (including Commission Services)	

Executive Summary:
Requirements specification based on the interviews

This document summarizes deliverable D1.2 of project FP7-284731 (UaESMC), a Specific Targeted Research Project supported by the 7th Framework Programme of the EC within the FET-Open (Future and Emerging Technologies) scheme. Full information on this project, including the contents of this deliverable, is available online at http://www.usable-security.eu.

This report is based on 24 semi-structured interviews with people from 6 different countries with the aim of finding usable applications for Secure Multiparty Computation (SMC). In addition, difficulties and challenges for the implementation were discussed. Based on the interviews we listed possible preconditions and barriers that need to be addressed for the implementation of SMC, and possible use-cases brought out by the interviewees.

The following preconditions were discussed in the interviews:

- There is a need for data sharing in more efficient ways.
- Participants are highly motivated to avoid data leaking.
- SMC can be a prerequisite if a trusted third-party organization is missing.
- The outcome of computation must be have value for the participants.

The following barrier categories were compiled based on the interviews:

- legal framework,
- communicating technology to the users and public image,
- risks deriving from the users,
- technology related barriers,
- data and computation related barriers.

The following use-case categories were compiled based on the cases brought out by the interviewees:

- (competing) entities comparing their data,
- different entities cooperating,
- pricing/price information,
- the state gathering statistical information,
- procurement and auctions,
- limited data usage inside the organizations.

List of Authors

Kadri Tõldsepp (UT)
Pille Pruulmann-Vengerfeldt (UT)
Peeter Laud (CYB)

Acknowledgements

We thank all the interviewees for their time and effort.

Contents

Introduction

This document is deliverable D1.2 of project FP7-284731 (UaESMC), a Specific Targeted Research Project supported by the 7th Framework Programme of the EC within the FET-Open (Future and Emerging Technologies) scheme. Full information on this project, including the contents of this deliverable, is available online at `http://www.usable-security.eu`.

This report contains the results of the interviews conducted with selected international experts from various fields. The first part of the report describes methodology and in the second part the results of the interviews are presented. In the first part of the results (paragraph 2.1 "Barriers") overview of the barriers expressed by the interviewees is given. The barriers are categorized in 5 different groups (legal framework, communicating technology to the users and public image, risk deriving from the users, technology related barriers, data and computation related barriers). Second part (paragraph 2.2 "Preconditions") describes preconditions that according to the interviewees should precede the implementation of SMC. In the third part (paragraph 2.3 "Use-cases") possible use-cases brought out by the interviewees are coded into 10 separate categories.

Chapter 1

Overview of methodology and presentation of interviewees

Methodology

Interviewees were contacted by e-mail that contained information about the project, description of the interview process and the Capability Model (deliverable D1.1). The aim of the Capability Model was to introduce the concept of SMC to parties not familiar with the concept. Capability Model includes short definition of terms connected to SMC and 12 illustrated example use-cases worked-out by the research team. Before the interviews the interviewees were asked to read the Capability Model and draw attention to cases that they can connect with their field of expertise.

Semi-structured interview design was used for the interviews. Interviews were conducted by 3 researches, all of them using the same interview design. The semi-structured interview design was complemented according to the interviewees' field of expertise (e.g relevant examples were used). Usage of the Capability Model was incorporated into the interview design. The cases illustrated in the Capability Model were used as a referral point in the interviews for the interviewees to find possible use-cases in their own field of expertise. For the interviewees who were already previously familiar with the concept of SMC, the Capability Model proved to be too general and therefore had a smaller role in the interview process.

For the analysis, the data were reduced by coding them into groups. The data on barriers were coded to 5 broad categories of barriers: legal framework, communicating technology to the users and public image, risk deriving from the users, technology related barriers, data and computation related barriers (see paragraph 2.1 "Barriers"). The first and second barrier category can be seen as external barriers: they are more related to the surrounding circumstances than to the technology itself. Other barrier categories can be seen as internal barriers, related directly to the technology or to the usage of the technology. Case numbers used within this report reference to the case numbers in the Capability Model (deliverable D1.1).

In addition to the barriers, interviewees also discussed several factors needed to be fulfilled for SMC to be implemented. These factors are shortly discussed in the paragraph 2.2 "Preconditions".

The data on use-cases were coded into 10 groups (see paragraph 2.3 "Use-cases"). The first and the second group are based on the motives of the participants: the first group lists cases were the participants are interested in comparing their data and the second group contains cases where the participants are interested in cooperation. The rest of the groups are more case based: one group contains descriptions of different versions of one case.

Cases enabled discussion on a fairly general level, which on the one hand yielded data that misses some details needed for future implementation, on the other hand, the general level enabled us to discuss sensitive data handling situations, which otherwise might not have been mentioned.

Interviewees

Interviewees were selected from various fields and countries. We conducted 24 interviews with experts from 6 countries. Interviews were conducted over period of April to June 2012. All interviews were conducted in a face to face setting.

Interviewee	Field of expertise	Organization	Country
I1	SMC specialist, informatics	University of Tartu	Estonia
I2	Security of computer networks	Estonian Information System's Authority	Estonia
I3	Ethics and philosophy of computers and computer based security		Estonia
I4	Securities operations and financial markets		Belgium
I5	Civilian registry	Local municipality government (Estonia)	Estonia
I6	IT-auditing, banking		Estonia
I7	Dairy industry	Estonian Dairy Association	Estonia
I8	Critical information infrastructure protection	Estonian Information System's Authority	Estonia
I9	Security of computer networks	Estonian Information System's Authority	Estonia
I10	Lawyer, legal systems		Estonia
I11	Biomedicine	University of Tartu	Estonia
I12	Cyber defense		Estonia
I13	Statistics, statistical offices	Statistics Estonia	Estonia
I14	Telecommunication	Cosmote	Greece
I15	Wholesale, agricultural markets		Estonia
I16	Marketing in retail sector		Estonia
I17	State defense specialist	Swedish Defense HQ	Sweden
I18	Infrastructure in healthcare sector	University of Skövde	Sweden
I19	Banking	Hanyang Cyber University, formerly worked in the Bank of Korea	South Korea
I20	Electricity markets	Nord Pool Spot	Estonia
I21	Product security management	Nokia	Finland
I22	Information security	Nokia	Finland
I23	Computer Science researher	Swedish Institute for Computer Science, formerly involved in Ericson Research Department	Sweden
I24	Tax offices, state registries		Sweden

In case of empty cells, the interviewee did not want to disclose the organization they represented. Full transcripts and audio files of the interviews are stored for future reference.

Chapter 2

Results

2.1 Preconditions

One of the purposes of the interviews was to find real-life use cases where SMC could be used. Therefore we discussed with the interviewees if they could see any potential for SMC in their field of expertise. In addition to discussing the possible use cases and challenges, the interviewees were also asked to explain why they thought there was/wasn't a need for something like SMC. Based on their answers we have compiled a list of preconditions for the implementation of SMC. On the one hand, these do have some overlaps with the barriers, and one could say that the barriers need to be overcome as a precondition for implementation, while on the other hand, this brings out new conditions for future implementation.

2.1.1 There is a need for data sharing in more efficient ways

Several interviewees argued that there is no need for functionalities that SMC is offering because as far as they know then the participants are content with existing solutions. For example I7 said that due to the smallness of the dairy market all the companies already have all the information they want to know and are allowed to know. The same was stated in the interview I13 and I23 for statistical offices: the interviewees didn't really see the necessity to implement SMC as, according to their opinion, it does not add a lot to existing solutions.

2.1.2 Participants are highly motivated to avoid data leaking

In the interview with I1, it was argued that SMC is only needed when the leakage of the data has major consequences. It was also argued that for most of the situation the SMC solution is too complex and offers security on a level that is not really necessary.

2.1.3 SMC can be a prerequisite if a trusted third-party organization is missing

I23 noted in the interview that in most of the situations it is possible to find a trusted third-party and that would also be simpler and more efficient. So alternatives are considered only for the problem/situation where it is not possible to involve a third party. Also, in the interview with the representative of Nord Pool Spot (I20), it was argued that using a third party to do necessary computation is sufficient. Nord Pool Spot acts as a neutral third-party on the electricity market. The interviewee said that clients do not really have a problem with that as they have trust in their organization.

2.1.4 The outcome of the computation must have value for the participants

I2 noted that probably the participants are not interested in implementing SMC if the outcome does not have high importance. He considered one of the possibilities to demonstrate the benefits of SMC by making computations that would provide unique data combinations. The gain needs to be obvious for smooth deployment.

Overall, many of the interviewees felt that SMC might be a good solution for them, if the benefits are obvious, if the trusted third-party solutions are not already in place, and that the motivations are there.

2.2 Barriers

During the interviews the challenges and difficulties of implementing SMC where discussed. In the following chapters we give an overview of the concerns and questions raised by the interviewees. The purpose of the analysis is not to discuss if the barriers raised by the interviewees are adequate. Rather we try to give a structured overview of possible concerns that should be addressed before turning to potential users of SMC.

2.2.1 Barrier category: legal framework

2.2.1.1 Limitations of current legal frameworks

Several interviewees brought out the limitation of current legal frameworks when discussing the implementation of SMC. Three of the interviewees mentioning legal issues were connected to state related organizations and therefore discussed the topic in connection to state related data handling. For example, according to I2 the current legal framework in Estonia lacks concepts that could adequately describe SMC. According to the I2, a current legal framework describes data anonymization or 'clean' data: neither of those concepts can be used to describe SMC adequately. I24 explained similarly that in the context of the Swedish legal framework, the Tax Office is either allowed to give data or not to give them: the law does not have the possibility of giving indications (e.g., a person's income is bigger than a certain amount). Interviewee 24 argued that several models in the Capability Model could be applied but not before the legal framework has been changed. The interviewee brought an example of gathering information from different data sources that would be interesting but that would mean that the current legal framework has to be amended. The interviewee stressed that this can be a long process, taking years. Therefore, the implementation of SMC in connection to the state related activities requires further analysis of the legal issues.

One interviewee (I1) brought out the legal issues also in connection to non-state participants: According to him, implementing SMC would require a new type of complex contracts between participants and, therefore, could bring additional costs.

2.2.1.2 Cartel and competition laws

When discussing the possible usages of SMC with I7 (representative of Estonian Dairy Association) it was repeatedly stressed that the cooperation between different entities in the dairy sector is restrained by cartel and competition laws. The interviewee remarked that probably there would be an interest among different entities in the dairy sector for some types of shared information but that the limitations set by the cartel and competition laws should be taken into account. According to I7, there is probably almost no information changed between the dairy companies, as that would be seen as criminal activity. The same issue was also brought out by I23 who remarked that SMC could be ill used to build cartels and for illegal information sharing. This information is especially relevant in connection with Cases 1 and 2 from the Capability Model.

2.2.1.3 Legal questions concerning possible use-cases concerning state databases (Case 7, 8, 11, 12)

Interviewees also raised concerns in regards to the cases using state databases. Currently it is legally not allowed to merge different state databases for any joint computation. In the cases mentioned above, SMC is used to make a query over several databases. I1 expressed concerns if the proposed solution is sufficient to remove the problematic issues due to which the legal restrictions concerning the merging of state databases were implemented. I2 raised the question of miners: according to him, currently it is legally not allowed to hand over state-owned data to non-state entities; therefore it could be an issue if the miners are private entrepreneurs even if the data are unreadable for them.

While, it is out of the limits of the current analysis to discuss if the possible legal issues raised by the interviewees are adequate, it can be said that the potential users are concerned with the legal issues and it is one of the factors that must be taken into account when introducing SMC to potential users.

2.2.2 Barrier category: communicating technology to the users and public image

2.2.2.1 Communicating the technology to the potential users in the private sector

Communicating SMC to the potential users was seen as one of the barriers in several interviews. I1 and I2 argued that the concept of SMC is too complex for non-experts to grasp. So selling SMC could be seen as selling belief of trustworthiness of the technology and should be marketed accordingly: it is important to realize what are the key aspects for non-expert users. The technological side may not have the highest importance for the non-expert users and should be therefore presented with adequate simplicity. It was argued that in order to sell SMC to the private sector, its trustworthiness should first be proven. One way of doing this is to have references from other users. The difficulty of communicating the technology was also reflected in many of the interviews. Several interviewees expressed interest in real-life cases where SMC has already been used.

2.2.2.2 Communicating the technology in the public sector (state databases, data provided by the citizens)

The question of communicating and marketing of the technology was also raised in connection to using SMC to state related data. The interviewees discussed, on the one hand, the importance of informing the decision makers and, on the other hand, the citizens. I8 brought out that the implementation of SMC in the public sector requires political support to begin with. He considered the implementation of SMC in the state sector as a good indication of support for the technology.

During an interview with one local government civilian registry official (I5) it was discussed how the public would receive this kind of new technology and data handling solution. For example, I8 said that the new technology/solution can be very good but in order for it to be successful it also has to be promoted to the public successfully. The interviewee brought examples of distrust among her clients and stressed the need to explain the public what is done and how it is done. In addition, the role of the media as an information channel was brought out. I2 brought out different aspects of informing the public: on one hand the public should be assured of the trustworthiness of the solution but they also should be informed how the data will be handled and used. Another interviewee (I10) mentioned that for him the most important question is the goal for what this technology will be used. Only after the appropriateness of the goal has been proven, the interviewee would look into questions concerning the technology itself (e.g., it's security).

2.2.2.3 Auditing and reliability/ trustworthiness

I3 discussed that it is important to demonstrate how the trustworthiness will be kept during the usage of the technology. So it is not only the question of showing that the technology in itself is trustworthy but that also everything that comes with it (e.g., users, procedures) can be relied on. The trustworthiness has to be demonstrated continuously, which means continuous auditing responsibility.

For different sectors and entities the importance of trustworthiness is different. It was argued that probably the question of auditing and trustworthiness has been given more attention in the public sector and in some parts of the private sector (banking). The question of trustworthiness was also discussed in the interview with the representative of Statistics Estonia. For this organization it is highly crucial to keep their reputation as a trustworthy organization as it is the cornerstone of their activities and success.

It is important to communicate the technology to the users and decision makers so that aspects relevant to the targeted group are taken into account and highlighted. Inadequate communication could become a barrier or have negative consequences. Trustworthiness will have to be one of the cornerstones of a successful approach: not only should it be demonstrated that the technology itself is trustworthy but also that the monitoring and auditing aspects have been thought about.

2.2.3 Barrier category: risks deriving from the users

2.2.3.1 User related risks

In several interviews the question of security was discussed in connection with users. The argument of the interviewees was that the technology in itself can be secure but it will not remove the highest security risk: the user. The opinion that the user is the highest security risk was directly brought out in 3 interviews (I3, I10, I21).

The concerns of the interviewees were different, though. One Interviewee (I3) described data handling as a chain where the weakest point normally is not technology. He considered there to be significant trust in encrypted channels and data handling procedures, but he considered other points of the data handling chain (e.g., people or processes) weaker and therefore more attractive for attacking. In an interview with a lawyer (I10) the possibility of intentional abuse by the users was discussed. The interviewee argued that the problem is not so much how the data are handled or what technologies are used, but how they can be protected against abuse from the users (in his examples, the interviewee focused mainly on unethical and illegal actions by the officials in state organizations). In interview I21, it was argued that malpractice does not so much occur due intentional abuse by the users but is more likely to be unintentional.

The interviewee I21 said that two important user related criteria are simplicity and usability. In connection with usability the interviewee mainly discussed the question of user involvement: he argued that this aspect should be thought through in the early stages of development. The user should not be involved too much (because it is demanding for the user) but involved at the right points (to still have some control). It is a question of finding the right balance and at the same time taking into account the variety of users that the technology can have. The interviewee argued that the so-called mainstream users cannot be expected to be as informed as the experts and, therefore, usage should be made simple and risk-free. In the same interview, it was discussed that the technology can be very simple but if the implementation is very complex and demanding, involving controls for the users, their user base will not be very wide.

2.2.3.2 Query formulation

Interviewee I3 and I23 stressed the importance of giving attention to the query formulation: queries should be only allowed to be formulated in a way where the outcome of the query does not lead to revealing information that must not be revealed. I3 discussed this issue in connection to case 5 in the Capability Model ("An organization interested in its members' info") and I23 discussed it in connection of combining the state databases. I23 argued that researches can formulate a query so that the outcome reveals more information than is allowed. When combining the state databases also the queries and the results of the queries have to be verified to guarantee that the privacy of the data providers is not violated.

2.2.3.3 Compatibility with 'honest but curious' model

As the state of the technology was understood at the moment, the idea of users following a pre-agreed protocol of 'honest, but curious' model seems to be crucial. There is apprehension that if the users will sway from the arrangements and make attempts to misuse the protocol, then data may still be lost. Interviewee I1 perceived that the trust model built in the technology assumes that the parties are behaving correctly themselves as the gain from multiparty computation is bigger than the gain from manipulating the data or process.

Potential users are quite likely to be concerned with different aspects of using the technology. Therefore the potential users are not only interested in the question 'how does it work?' but they also need to be aware of the user-associated risks that technology does or does not handle. Additionally, they are interested in the extent of the potential damage caused by inadvertent or intentional misuse of the technology, processes, and data handling.

2.2.4 Barrier category: technology related barriers

2.2.4.1 Miners

The question of miners was discussed indirectly in two cases: two of the interviewees stated that the data handling should physically remain within the borders of the state. I8 said that it is important for all the information related to the state and the state security to remain physically within the borders of the state. The same factor was brought out in the interview with I13 (statistics office): the interviewee said that they used services located outside the country borders for testing their information system but it will not be possible to do with real data. Therefore it can be concluded that at least with certain data all the miners should be located within the country. I11 argued that this could be problematic. The interviewee was not convinced that it is possible to find at least in Estonia 3 suitable miners. Those 3 miners should of course technologically be SMC capable but it should be also guaranteed that there are no possibilities for them to start co-operating.

2.2.4.2 Scalability

In three interviews (I2, I3, I21) the interviewees discussed shortly the question of scalability as a necessary factor/attribute. Interviewee I2 said for wider breakthrough of the technology, the scalability needs good and compatible references (e.g., the fact that the technology has been implemented in Estonia may not serve as a sufficient reference to be implemented in Germany).

The interviewees had some concerns with the technology itself, mostly with the issues of controllability of data handling and capacity and scalability of the technology.

2.2.5 Barrier category: data and computation related barriers

2.2.5.1 Necessity to 'see' the data

Several interviewees were hesitant about the idea of doing computation without having the full data available. Interviewee I11 argued that it is necessary to have a full overview of the data when doing scientific analysis. According to the interview not seeing the data makes analyzing it for the scientist more complicated. So scientists can be skeptical of a solution that does not allow them to see all the data. It is also possible that the quality of the analysis suffers.

Interviewee I13 said that statistical offices need to have full overview of the data. According to the interviewee, full overview of the data is needed to guarantee the quality of data. Another reason is that when combining different databases, the databases need be described according to set standards (interviewee brought an example of SDMX (*Statistical Data and Metadata eXchange*)). In order to describe the databases and to combine them there has to be at least one side that has a full overview.

The necessity to see the data was also brought out in connection with the electrical exchange market (I20). The representative of an organization that manages the electrical exchange market in different countries expressed the need to have a full overview of their clients' data in order to manage the exchange successfully. As explained by the interviewee, the exchange involves more complicated algorithms than just calculation of the market price.

So far, many existing data handling procedures have used the possibility to "see" the data. For usable applications, SMC needs to investigate further the possibilities to replicate these processes without the full disclosure of the database and communicate this to the potential user.

2.3 Use-cases

One of the aims of the interviews was to find real life examples where SMC could be implemented. The interviewees were asked if they see problems and situation in their field of expertise or anywhere else that could be addressed by SMC. The details of possible cases were discussed with the interviewees but some of the cases identified are quite general as it was not always possible for the interviewees to give very detailed analysis. Below we have listed all the cases identified by the interviewees. The data on use-cases were coded into 10 groups. The first and the second group are based on the motives of the participants. Other groups are more case based: one group contains descriptions of different versions of one case. Note that the purpose of this analysis is not to discuss if and how problems and cases raised by the interviewees can be solved by SMC.

2.3.1 (Competing) entities comparing their data

2.3.1.1 Case: universities comparing their data (discussed in interviews I1, I2, I13, I16)

Participants: different universities, Tax Office
Outcome of computations: statistics on revenues of graduates of different universities
Description: This case can be categorized under the Case 1 (Three competitors) or Case 11 (State database and interested non-state parties) in the Capability Model. I1 said that universities could be seen as competing companies who are interested in each other's data. In the interview with I1 it was discussed that competing universities could compare "value" of their education by analyzing their graduates' net revenues that the Tax Office can provide. However the interviewee was not sure what would be the motivation of the Tax Office to participate.

2.3.1.2 Case: quality of services in public and private sector (I3)

Participants: any public or private organizations offering some kind of services
Outcome of computations: comparable data on quality of the services
Description: The interviewee proposed that the quality of services could be compared in several sectors where it currently is not done so frequently. The interviewee brought out the concrete example of the IT sector and different organizations in the public sector (for example Citizenship and Migration Bureau). The aim of comparing the data is to enhance the quality of services offered.

2.3.1.3 Case: losses due to fraud/illegal activities in banking (I6, I12)

Participants: banks, state organizations
Outcome of computations: comparable data on losses due to fraud/illegal activities
Description: According to the interviewee I12 banks are quite innovative when it comes to cooperation and implementing new technologies. According to him, currently the banks do not share information about the losses they have suffered due fraud of illegal activities. However this information would be interesting for the banks and also for some state institutions. The same problem was discussed in interview with I6. According to him, this information is actually being shared between participants in Estonia. However he did see the potential in sharing information about losses on a bigger scale (for example between banks in the whole of Europe). The interviewee also mentioned that comparing these kinds of data could be complicated as it is actually quite difficult to calculate what the exact losses are and the methodology may vary considerably.

2.3.1.4 Case: data comparisons in the retail sector (I16)

Participants: retail companies
Outcome of computations: different rankings (e.g., amount of visitors)

Description: The interviewee said that it would be interesting for the retail companies to compare the number of visitors but none of the companies are interested in sharing their data. Also comparing service quality in a privacy preserving way would be interesting for the companies.

2.3.1.5 Case: data comparisons in telecommunications sector (I22)

Participants: telecommunication companies
Outcome of computations: different data comparisons (e.g., salaries, sales)
Description: According to the interviewee, SMC could be a possible solution for telecommunication companies to do different data comparisons. The interviewee brought concrete examples of salaries of the employees and sale numbers. Although these data are also shared now (salaries are compared through a third-party organization), it would be interesting for the participants to use SMC if it lowered their costs and brought more efficiency. The interviewee also introduces the concept of "benchmarking": for example it is interesting for the companies to compare certain cost factors with those of other companies. However, the interviewee said that currently this information is absent from the market.

2.3.2 Different entities cooperating

2.3.2.1 Case: electricity markets cooperating (I20)

Participants: different electricity markets
Outcome of computations: for example, price coupling
Description: European electricity markets are developing cooperation projects (the interviewee mentioned a project organized by the European Commission) and as part of this cooperation they have to exchange information. However at the same time they also have to keep their clients' privacy.

2.3.2.2 Case: planning agricultural production (I15)

Participants: agriculturalists
Outcome of computations: overview of planned agricultural products
Description: The interviewee said that it would be useful to have an overview of what kind of agricultural products different producers are planning to grow. This kind of overview would help to avoid overproducing certain products. At the same time, it would also show if there is a lack of some products. As there is no such overview now, it can happen that supply of some products is bigger than the demand.

2.3.2.3 Case: information sharing by military or intelligence agencies (I12, I17)

Participants: military or intelligence agencies of different states
Outcome of computations: shared information
Description: Interviewee I14 explained that sharing critical information between different military or intelligence agencies is problematic. He brought an example of situational awareness: different entities can have different information and they are not interested in sharing all the information but only relevant parts of it. Similar problems were raised by interviewee I17. This interviewee concentrated on problems of exchanging information between different state organizations. On the one hand it is difficult to exchange certain information due legal issues. Another issue is that different entities are competing and are not interested in sharing all the information they have. It is also important to keep the privacy of the information sources.

2.3.2.4 Case: information exchange between different state organizations (I24)

Participants: different state organizations
Outcome of computations: information exchange
Description: I24 described that there is a need to exchange information between different state entities and this could be done in more efficient ways. The same possible use-case was also mentioned in the interview

with I23, who said that SMC could be used for cooperation between different state organizations that don?t have full trust in each other

2.3.3 Pricing / price information

2.3.3.1 Case: pricing in the retail sector (I16)

Participants: different retail companies or any other companies concentrating on selling
Outcome of computations: price comparisons and rankings of companies
Description: The interviewee thought that retail companies would be interested in comparing their prices with other companies. Although the information about the prices is publicly available, it requires too much work-force to do the comparisons. Therefore the companies could be interested in a system that would give their ranking in connection to a price of a specific product. The interviewee remarked that the same would apply to small companies in other sectors (e.g., gardening) who would like to compare their prices with other companies.

2.3.3.2 Case: pricing in marketing (I16)

Participants: marketing companies and different companies/entities buying marketing services
Outcome of computations: price overview of different marketing services
Description: The interviewee said that it is very difficult to know if the prices offered by the marketing companies for their services are adequate. The price negotiations are long and the discounts achieved can be considerable. Therefore there is a need to know what the average prices for marketing services are. However, the interviewee said that it is doubtful if participants are willing to share this information.

2.3.4 The state gathering statistical information

2.3.4.1 Case: data collection in dairy industry (I7)

Participants: different dairy companies
Outcome of computations: statistical information needed by the Ministry of Agriculture
Description: According to the interviewee, there is an issue with gathering statistical information for The Ministry of Agriculture. Apparently The Ministry of Agriculture is interested in the price margins but does not have this information. The interviewee said that this information is known among the dairy companies. Possibly there would be a need for a more efficient solution that would help also The Ministry of Agriculture to get the wanted information in a privacy preserving way.

2.3.4.2 Case: black market and reporting on corruption (I12)

Participants: individuals wishing to report in a privacy preserving way, state organizations gathering statistical information about the size of the black market, the level of corruption
Outcome of computations: statistical information
Description: The interviewee suggests that there might be a need for a solution that allows individuals to report delicate information in a way that their privacy is protected. SMC could for example be used to collect information about corruption: individuals may be reluctant to give out this information if their identities are connected with the information.

2.3.4.3 Case: collecting information on cyber security incidents (I2)

Participants: the state organization responsible for the management of security incidents in computer networks, different internet service providers
Outcome of computations: overview of security incidents

Description: If a client of some ISP has malware, and the activities of this malware are visible on the Internet, then the activities may originate from different IP addresses during different times. The ISP knows which IP addresses at which times correspond to the same client but they are not allowed to release that information. Organizations dealing with security incidents would like to have an overview of malicious activities but they do not need to have the identities of the users. Ratings could be used to evaluate the behavior of the customers of internet service providers and the identity of the customer is only revealed when certain level has been exceeded.

2.3.5 Combining state databases

Participants: state databases and interested parties
Outcome of computations: information/statistics from state databases
Description: This case was discussed in several interviews (I1, I5, I8, I11, I13). It was agreed that combining different state databases can have important practical value and keeping the privacy of the data providers is crucial. Mostly combining databases was discussed in the context of doing scientific researches. I11 argued that they frequently need to analyze patients' data and need to connect different databases. Currently they can only link different databases if they have the whole information (e.g., identity of certain patients) from one database and then turn with that information to another database. However that specific information is not needed for the statistical analysis and is therefore an unnecessary security risk.

2.3.6 Procurement and auctions

2.3.6.1 Case: state organizing procurements and auctions (I3)

Participants: state organization organizing procurements and auctions
Outcome of computations: -
Description: The interviewee thought that SMC could be useful in connection to state procurements and auction as there are some problematic issues and better solutions are sought but it was not discussed what role SMC could play.

2.3.6.2 Case: auction environment for agricultural products (I15)

Participants: sellers and buyers of agricultural products
Outcome of computations: optimal prices, price curves
Description: According to the interviewee there would be a need for an auction environment for agricultural products that helps the farmers sell their products with optimal prices. Currently it is difficult for the farmers to know what price they should ask for their products, as they do not have that information. The auction environment should be confidential, so SMC could be used to guarantee the confidentiality without involving third parties.

2.3.6.3 Case: modified Dutch auctions and Dutch auctions for securities (I4)

Participants: security owners wishing to sell their securities back to the issuer
Outcome of computations: security owners have overview of (aggregated) bid prices during the process of the auction
Description: To buy back securities issued by them, the issuers of the securities may organize Dutch auctions or modified Dutch auctions. The procedure of those auctions involves bid-prices that have to be offered by the security holders. During the auction (which may last for some weeks), only the issuer has full overview of the bid-prices offered. However it would be also interesting for the holders of the securities to have that overview. The issuers are in that respect not neutral information sources. At the same time, it is likely that the security holders want to keep their bids secret.

2.3.7 Limited data usage inside the organizations

2.3.7.1 Case: limiting data usage for the officials in the public sector (I5)

Participants: state organizations
Outcome of computations: officials have access only to the data needed for their tasks
Description: The interviewee said the possibilities for the officials to misuse the data must be limited as much as possible. Therefore, officials should not have access to data they do not directly need to fulfill their tasks. So for example officials should not be able to see each record if they only need to have the general numbers.

2.3.7.2 Case: limited internal data usage in banking (I6)

Participants: banks, different departments in banks
Outcome of computations: limited data usage inside the banks
Description: Different departments in the bank need to use different information about the client and should only have limited access to client information.

2.3.7.3 Case: patient data usage by healthcare sector workers (I18)

Participants: different entities/organizations in the healthcare sector
Outcome of computations: -
Description: The interviewee described an issue connected to patient data usage. All the information concerning the patient is treated as an entity and is not categorized as separate events. Therefore, different healthcare organizations have access to the whole information, including also the information they do not need.

2.3.8 Market research/analyzing client data

2.3.8.1 Case: market research companies combining databases (I1, I8)

Participants: different market research companies
Outcome of computations: statistics based on combined databases
Description: It was suggested by two interviewees that SMC could be used to conduct market researches. However, they were not experts in that field. I1 described how different market research companies could sell information gained from combining their databases. The same use-case was described in the Capability Model under Case 9.

2.3.8.2 Case: banks analyzing client data (I6)

Participants: banks
Outcome of computations: information based on client data
Description: According to I6 the information that banks have about their client is currently not analyzed very actively. He proposed that if this information could be analyzed in a way that the privacy of the clients is protected then this information could be shared between the banks and used also outside the banks (e.g., to analyze the consuming habits of the whole population). I6 also argued that if more information connected to clients is analyzed and brought to the public then it could make the market to work more efficiently.

2.3.8.3 Case: telecommunication companies analyzing client data (I22)

Participants: telecommunication companies
Outcome of computations: information based on client data

Description: I22 said that the telecommunication companies are already analyzing client data and using it, among other things, for marketing, but as it is very critical information then they are constantly looking for new solution that can guarantee more security but at the same time keep the costs down.

2.3.9 Reporting

2.3.9.1 Case: accountability in banking (I6)

Participants: banks, organization banks report to
Outcome of computations: reports
Description: Banks are obligated to report on their activities to several controlling institutions. For example, in Estonia the banks have to report information concerning their clients (movement of customers of pension funds) to the Financial Supervision Authority. According to the interviewee I6, SMC could be used for the reporting.

2.3.9.2 Case: accountability in telecommunications (I22)

Participants: telecommunication companies, the organization that companies report to
Outcome of computations: reports
Description: The interviewee explained that telecommunication companies are annually disclosing information about security implementation to the International Security Forum. According to the interviewee, this organization is a trusted third party, but SMC could be an alternative solution.

2.3.10 Verification

2.3.10.1 Case: verification in corporate actions (I4)

Participants: security owners and their representatives, security issuers and their representatives, entities organizing corporate events
Outcome of computations: security owners or their representatives can confirm accordance with the conditions of the corporate auction in a privacy preserving way
Description: Often several legal restrictions apply to corporate events. For example it may be required that security holders from a certain country may not participate in the organized corporate event. Currently it is difficult to check if those requirements are always fulfilled because issuers and their representatives do not know the identities of their security holders. Therefore all responsibility is left on the security holders who have to decide themselves if they can participate in the corporate action or not. The interviewee argued that it would be interesting to have a solution that allows verifying the rightfulness of security holders to participate in the corporate auction without really knowing the identity of the security holder.

Chapter 3

Detected privacy-sensitive problems

We have identified a set of privacy-sensitive problems that UaESMC will address. We have abstracted these problems from the concerns mentioned in several interviews. In the subsequent work in UaESMC, we plan to formalize them into algorithmic problems that can be solved with SMC. We will now describe each of the identified problems.

Input verification This topic arose in the context of financial entities submitting their bids to auctions. The problem is to verify that the input submitted by an entity conforms to a set of rules. We believe this issue is relevant in other applications as well. This leads to (anonymous) credentials and efficiently, but privacy-preservingly checking whether the input matches the constraints listed in the credential.

Database operations Several of the interviewees were considering the applications that would be possible if one could refer to several personalized databases at the same time. The computations themselves would likely be rather straightforward database queries, consisting of joins (over private columns), simple filterings (equality, greater/less than) and aggregations (sum, maximum over rows satisfying certain conditions). Privacy-preserving database engines are currently still in their infancy and it makes sense to research this area in UaESMC. In particular, it may make sense to try to optimize the query strategies taking into account the relative efficiencies of privacy-preserving operations.

Outlier detection Interviewees working in the fields of security may want to see whether something is wrong, and this can be seen only (or better) if several databases are combined. The particular algorithms will probably depend on the area of activity, but the outlier detection algorithms from the data mining community is the first step here.

Optimization problems Several optimization problems emerged from a couple of interviews. In one of the problems, one data provider had the description of a landscape with secret artifacts on it, and another data provider wanted to navigate this landscape, going from point A to point B (again secret). This directly leads to a shortest path problem in a suitably defined graph. CYB has already started investigating this problem (deliverable D2.2.1, due in January 2013, will describe our results). The developed methods look promising to be applicable to linear programming problems in general, e.g., the planning of agricultural produce (what to plant, how much and where). Any further, more complex constraints on optimization problems may easily turn them NP-hard, hence it also makes sense to study privacy-preserving methods for solving NP-hard optimization problems. Genetic programming may be particularly simple to adapt for that purpose.

Computational geometry Navigating a landscape, as mentioned in the previous paragraph, also has elements of computational geometry in it. We will investigate if SMC can be used to reduce the search problem in some geometric structure to a search problem in an abstract graph.

Incentives for participation The interviews also demonstrated computations into which the data providers would be reluctant to input correct data not because of privacy issues, but because they would not like the results of that computation. One example was finding out the premiums earned by different parties in a supply chain (from farmer to grocery store). The government might be interested in learning the average premiums at each step. On the other hand, the parties might prefer even not the averages to become "verified" public knowledge, as it might give negative publicity to their sector. Note that while the computation itself would be rather simple in this case (subtracting prices, and averaging), the game-theoretic aspects are much more complex and significant. These aspects, in turn, can lead to more complex SMC solutions.