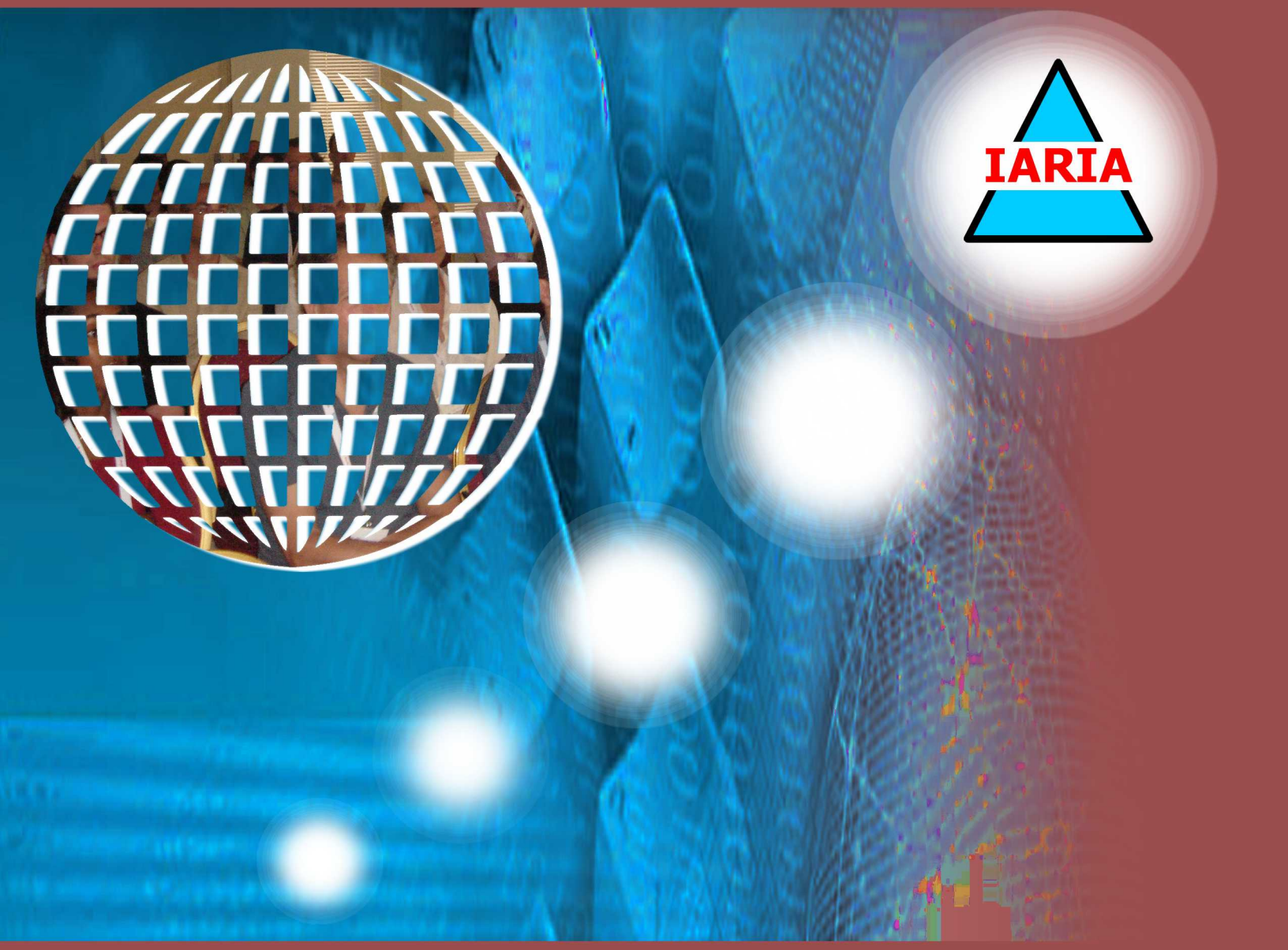
International Journal on Advances in Security



2010 vol. 3 nr. 3&4

The *International Journal on Advances in Security* is published by IARIA.
ISSN: 1942-2636
journals site: http://www.iariajournals.org
contact: petre@iaria.org

Sponsored by IARIA
www.iaria.org

International Journal on Advances in Security
Volume 3, Number 3 & 4, 2010

- Aljosa Pasic, ATOS Origin, Spain
- Vladimir Stantchev, Berlin Institute of Technology, Germany
- Michiaki Tatsubori, IBM Research - Tokyo Research Laboratory, Japan
- Ian Troxel, SEAKR Engineering, Inc., USA
- Hans P. Zima, Jet Propulsion Laboratory/California Institute of Technology - Pasadena, USA // University of Vienna, Austria

Security in Internet

- Evangelos Kranakis, Carleton University, Canada
- Clement Leung, Victoria University - Melbourne, Australia
- Sjouke Mauw, University of Luxembourg, Luxembourg
- Yong Man Ro, Information and Communication University - Daejon, South Korea

International Journal on Advances in Security
Volume 3, Numbers 3 & 4, 2010

CONTENTS

Enhancing Law Modeling and Analysis: using BPR-Based and Goal-Oriented Frameworks

Komminist Weldemariam, Adolfo Villafiorita, Angelo Susi
Center for Information Technology
Foundation Bruno Kessler (FBK)
Via Sommarive 18, Trento, 38123 Italy
Email: {sisai,adolfo,susi}@fbk.eu

Alberto Siena
Faculty of Computer Science
Free University of Bolzano
Bolzano, Italy.
Email: alberto.siena@unibz.it

***Abstract*—Legal documents contain regulations and principles at different levels of abstraction. They constitute rich sources of information for public administrations (PA) redesign and eventually for the software delivery that must comply with normative regulations that are specified in laws and procedures. In order to facilitate the alignment between these elements, systematic methods and tools automating regulations modeling and analysis must be developed. In this paper, we propose the integration of process modeling (named VLPM) and goal-oriented (named Nòmos) tool-supported methodologies to systematically model and analyze laws and procedures in public administration. We show that such integrated view would provide a framework that allows tracing and reasoning either top-down, from the principles to the implementation or, vice versa, bottom-up, from a change in the procedure to the principles. Finally, we also believe that this would provide a facility for interchanging models among different tools and for sharing models among different actors.**

***Keywords*—BPR; goal-orientation; laws & procedures; Nòmos; public administrations; regulation compliance; VLPM.**

I. Introduction

The introduction of new systems and procedures requires a careful modeling and analysis of laws to ensure that no conflict arises between the way things are done (i.e., processes) and the way things are meant. In that context, this paper provides our efforts in modeling and analysis of laws and procedures, which has its roots in process modeling and goal-oriented frameworks. This work is based on a conference paper on the International Conference on Technical and Legal Aspects of the e-Society 2010 (CYBERLAWS'10) [1].

Typically, there are three elements on which governments can operate to improve their public administrations (PA). One important contributor to the efficiency and improvement of PAs is the way in which the processes is (re)designed and developed. The use of Business Process Re-engineering (BPR) in that respect has become one of the recent trends undertaken to redesign processes, reduce costs and improve citizens' participation in favor of PAs [2], [3], [4]. However, there is a need to link procedures to the regulations by which procedures are defined and directed within legal documents, which contain information vary with respect to the levels of abstraction. They also describe principles or general rules that have to be followed, and thus requiring the implementation of related processes.

Another important fact to mention is the social relevance of information systems, which is determined by the way the information system is initially conceived. If misaligned with legal prescriptions, a functionality of the system can violate the rights of users, thereby breaching the law [5]. Note, however, that the system itself is not responsible for the breach, but rather, it is the owner, the designer and/or the operator who are responsible for the breach. Preventing this situation to happen is in the hand of those who are called to define the system's functionalities: the requirements analysts.

In fact, there are various broad approaches that can be used to mitigate (part of) the mentioned challenges, e.g., see in [6], [7], [8], [9], [10]. Different but complimentary approaches are often preferable for different types of challenges within a domain, and a combination of approaches would sometimes be desirable. More specifically, the integration of tool-supported methodologies with the aim of supporting the different levels of abstraction in PA processes can make easier the modeling and analysis of laws and procedures. One way to do this, for instance, is by using tools and techniques to model and analyze the underline low-level concepts of the laws as a business process. With different approach, move the emphasis of the modeling and analysis of the principles and procedures to a higher-level abstraction by interpreting each individual piece of information extracted from the law or principles as a root goal that can be decomposed into one or more subgoals. The results of these approaches would then be assessed, refined, and integrated systematically.

Along this direction, previously we presented an approach that takes the advantages of two existing tool-supported methodologies to assist the different aspects of law modeling and analysis. The first approach is based on goal-oriented framework —named Nòmos [11], [5]. The second approach, whereas uses the notion of process modeling based on subset of UML diagrams —named VLPM [12], [13]. Both approaches offer related tools to support their methodology and to allow traceability between the law and the corresponding models at different level of abstractions. We also showed that they individually have significant limitations. For example, the VLPM does not provide notations and means to represent the principles behind the procedure and to reason about possible alternative implementation; instead, Nòmos does not provide

low-level implementation of the actual processes impeding the analysis of some components of the law.

This paper extends the work presented in [1] by further elaborating the context of the two approaches. We detail the underlying meta concepts and the current improvement of the two frameworks individually and as combined in handling (some of) the peculiarities in modeling and analysis of laws. For example, the VLPM includes semantic knowledge through the use of ontology. Its model sharing is now general enough since all the information can easily be stored as RDF statements [14] to maintain links between parts of the documents, parts of the process models and also elements of other types of models. It is exactly this connection that adds value to the solution we propose and makes our approach more significant than the simple juxtaposition of the two techniques. We also provide a proof of concept of the advantages we can get by putting the two together with an example.

The paper is organized as follows. The next section discusses the BRP context of modeling and analysis of PAs in connection with laws. Section III discusses the goal-oriented approach for law compliance by detailing how such processes using the Nòmos framework can be realized at higher-level. Section IVdiscusses the VLPM approach for supporting PAs with new extensions. We discuss our attempt to support BPR using the best out-of these two techniques with example in Section V. Related work, and conclusion and possible future directions are provided in Sections VI and VII.

II. Motivation

In recent years, new laws have been enacted to explicitly regulate sensitive information related to businesses and healthcare. Existing laws have been revised and also gained new meaning when referred to an Internet-based activity. As a result of this, concerns like security, privacy and governance are increasingly the focus of (digital) government regulations around the world. This trend has also created challenges in the definition and redesign of public administration (PA) processes and in the compliance of regulations. Consequently, different entities —such as PA officers, lawmakers, software engineers— are required to ensure that their software delivery complies with relevant regulations, either through (re)design or (re)engineering activity of a particular project.

In fact a decision in (any) project preliminary phase has more relevant effects than those delayed to the subsequent ones [15]. In the same way, normative choices and changes in PA influence the law effective applicability with respect to the desired system. In principle, we distinguish three different kinds of reengineering projects:

1) *System automation level.* The goal of this kind of project is introducing a new system to better support one elementary task or limited procedure. Typically small in scope, these kinds of projects provide limited improvements but are simple to implement, since they do not affect neither the procedures nor the laws.
2) *Departmental level.* The goal of this kind of project is changing the way in which work is performed within a functional unit, (often) to make it more rational and efficient. These kinds of projects are more impacting, as they require some kind of re-organization of the work, often accompanied by the introduction of new ICT systems. The impact on the laws, however, is null or minimal.
3) *Inter-departmental level.* The goal of this kind of project is providing a better implementation of those processes that involve different departments or possibly change the allocation of responsibilities or both. It is the case, for instance, of decentralization projects, where competences are moved from central government to districts. These kinds of projects are clearly the most impacting, since they act at all levels, including the normative one.

The first kind of project is a standard software development project for which there is a rich choice of tools, development cycles, and project implementation alternatives. However, the other two kinds of projects pose two peculiar and closely related challenges, which root is in the relationship between the laws and the procedures that implement such laws. These challenges are particularly common in PA domain, which are

- *Laws provide the framework that constrains and limits its possible choices and alternatives in reengineering processes*. Providing tools and notations can allow to explicitly model and reason about the alternatives and constraints, and as the same time could help to develop more efficient solutions.
- *Laws and processes are intertwined as requirements and implementation are in software development processes.* Providing tools to explicitly trace the connection between laws and process elements helps for a more efficient and coherent management of the system. This can help ensuring that procedures correctly implement the law, and at the same time it could help to understand which laws might be affected by a change in the processes.

Interventions usually require to change part of the law. However, in order to understand where and how to modify the law we have to set, prepare, and validate new processes as well as to recognize the new roles and people responsibilities. Additionally, regulating a complex system or a new one requires to understand about the procedures to activate in order to answer questions like who is in charge of, when the task should be done, how to face exceptional situations, and so on. These all call a significant reform needed to provide correct snapshot about the existing procedures, to propose the (re)design and development of a new system.

As said before, one of the tools to enact this reform is the application of BPR techniques. In PA, this is an activity which involves (independently or in collaboration) law-makers who amend laws, process designers who try to optimize existing processes, and software developers, to support existing processes and/or procedures with technology. Modeling facilitates the communication and understanding of the actual organization among these users and is helpful in building a shared vision between domain experts and technicians. It also

provides an easier way of analysis in order to evolve towards efficient and higher quality processes, if not pose related risks.

Unfortunately, most of the existing modeling techniques were developed with the aim of optimizing supply chains or production. This indicates that there are no as such clear goals in modeling public workflows [4]. Thus, on one hand, it is essential to know the constraints established by the law, but, on the other hand, it is also necessary to support the office in charge of such processes in order to decide the changes required by the new processes. However, one of the major difficulties encountered in this domain is the strong dependency between processes and laws. As said before, any implementation of software delivery requires a parallel action on both the redesigning of processes and on the introduction of law changes. This means that the current law (in a sense: rather than the processes), must be considered as the constraint, the engine, and the target of the reengineering activity. This link between models and laws raises further issues related to the maintainability of the models over time, since it is necessary to guarantee coherence of the models with the laws in order to have the models retain their value. We argue that these situations can reasonably be tackled by providing homogeneous and structured models of the current processes (the business architecture), which in turn should allow to redesign the new software delivery.

As we hinted in [1], it is important to devise a methodology and tool that should help tackling the two aforementioned challenges. In fact, we can apply techniques (e.g., goal-oriented methodology) that can help tackling the first problem by providing precise notations and alternatives to avoid misinterpretation and resolve ambiguities that can arise, and by performing high-level formal reasoning. The Nòmos framework [11], [5] fits for this purpose. In contrast, the second challenge can be tackled through a proper BPR approach —namely, by devising process modeling and redesigning methodology and by developing its supporting tool. Thus, the VLPM —a tool we developed for the purpose, by extending work [16]— helps tackling the second challenge.

III. Goal-oriented Approach for Law Compliance

When facing law we need to know the concepts used by law to give prescriptions. Law is grounded on the notion of ***right***, which can be defined as entitlement (not) to perform certain actions or be in certain states, or entitlement that others (not) perform certain actions or be in certain states [17]. Rights are classified by Hohfeld in the 8 elementary concepts of *privilege*, *claim*, *power*, *immunity*, *no-claim*, *duty*, *liability*, *disability*, and organised in opposites and correlatives (see Table I).

TABLE I
The Hohfeldian taxonomy.

Legal relation	Opposite	Correlative
Claim	Noclaim	Duty
Privilege	Duty	Noclaim
Power	Disability	Liability
Immunity	Liability	Disability

Notice that in commonsense we might call *right* as a duty or a liability, since the word has a slightly different meaning. Here the more intuitive meaning of "right" is a **Claim**, which is the entitlement for a person to have something done from another person who has therefore a **Duty** of doing it. For example, if John has the claim to exclusively use of his land, others have a corresponding duty of non-interference. **Privilege** (or liberty) is the entitlement for a person to discretionally perform an action regardless of the will of others who may not claim him to perform that action, and have therefore a **No-claim**. For example, giving a tip at the restaurant is a liberty, and the waiter cannot claim it. **Power** is the (legal) capability to produce changes in the legal system towards another subject who has the corresponding **Liability**. Examples of legal powers include the power to contract and the power to marry. **Immunity** is the right of being kept untouched from other performing an action, who has therefore a **Disability**. For example, one may be immune from prosecution as a result of signing a contract.

Two rights are **correlatives** [18] if the right of a person A implies that there exists another person B (A's counterparty), who has the correlative right. For example, if someone has the claim to access some data, then somebody else will have the duty of providing that data, so duty and claim are correlatives. Similarly, privilege-noclaim, power-liability, immunity-disability are correlatives. The concept of *correlativeness* implies that rights have a **relational nature**. In fact, they involve two subjects: the owner of the right and the one against whom the right is held —the *counterparty*. Vice versa, the concept of *opposition* means that the existence of a right excludes its opposite.

The choice to adopt the Hohfeldian taxonomy of rights is due to several factors. First, as said above, the importance it has in the juridical literature suggests that this is actually the kind of information that we need to know about law. Second, Hohfeldian has a range of concepts and level of abstraction. These make its representation capabilities very close to the expressiveness of legal texts. In fact this consideration mainly comes from experience: constructs like powers, immunities and so on do actually exist in legal texts. Differently, for example, from deontic logic-based approaches, the proposed taxonomy is able to successfully capture them. Finally, the Hohfeldian concepts that have a *descriptive* nature, rather than *prescriptive*, acting as the bridge between the world of the "ought", typical of legal prescriptions, and the world of domain description. However, the Hohfeldian concepts do not prescribe what stakeholders should do, but rather, they describe what are the legal relations that bind them. This is of particular importance in requirements engineering, whose first step (early requirements analysis) is to describe the so-called "as-is", before specifying the "to-be". So, linking a description of stakeholders' goals with a description of applicable laws can allow to reason about compliance and compliant alternatives. In the following, we (formally) characterize such a link.

Rights are not symmetric: the position of the *claim* owner is different from the position of the *duty* owner. Generally speaking, the two positions are called active (juridical) position and

passive (juridical) position, and each right has exactly an active position and a passive position. To capture this characteristic in the meta-model, active `Actors` are in `holder` relation, while passive `Actors` are in `counterparty` relation with respect to the right (see Figure 1).

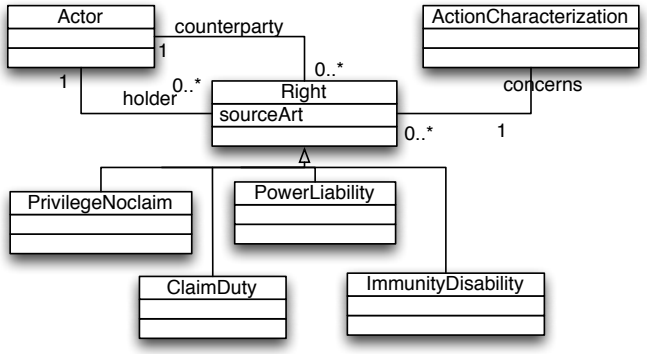


Fig. 1. The Nòmos metamodel: elements of a normative proposition.

The last component of a normative proposition is called *action*, the actual object of the right. An "action" designates a description of the set of admissible states of the world. To avoid confusion with a more common use of such word, we refer to it as `ActionCharacterization`. Each `Right` is in `concerns` relations with exactly one `ActionCharacterization`, but an `ActionCharacterization` can be addressed by a number of rights, as depicted in Figure 1.

A. Modeling the Structure of Law

The concept of normative proposition allows to split the complexity of legal statements into atomic elements. But the legal prescriptions contained in laws have more properties that have to be considered. In particular, legal prescriptions are articulated structures built with conditions, exceptions, and so on. It is important to capture the effects of these conditionals in order to obtain a meaningful requirements set. We give a uniform representation of conditional elements by establishing an order between normative propositions. For example, a citizen may have the duty to give his personal details to the policemen. However, if the policeman does not identify itself correctly as a policeman, then the citizen is free whether to do it or not. Instead of trying to formalize the *if [...] then [...]* condition, we split the problem in three steps. First, we define a first right, $r1$, —a duty of the citizen —to give personal details to the policemen. Second, we define another right, $r2$, —a privilege of the citizen —to give personal details to other citizens. Finally, we establish an order between the two: $r1 > r2$ which means that, whenever $r1$ is applicable, $r2$ is not. This is captured in the meta-model of Nòmos by the concept of **dominance** (class `Dominance`), as shown in Figure 2. This is connected to the class `Right`, which establishes the priority of the *source* right over the *target* one.

The normative propositions, manually extracted and ordered according to the meta-model, are put together to form the

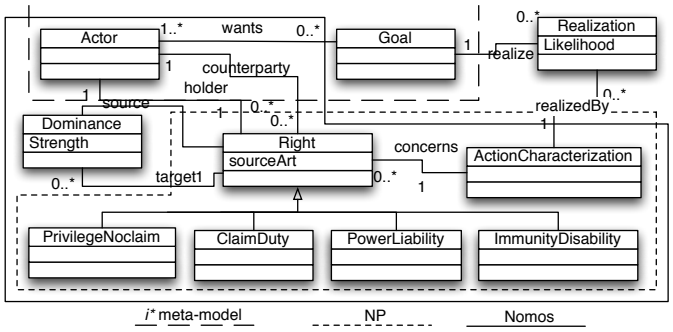


Fig. 2. The link between rights and goals as proposed by the Nòmos meta-model.

model of the law. As the meta-model shows, such a model does not contain anymore information on the physical structure of the law given by its nesting into titles, paragraphs and so on. Moreover, it it does not contain cross-reference information. However, to each right (class `Right`) carried by a normative proposition, we are able to associate its source article or any further information to precisely record where does the normative proposition come from. This would allow to ensure full traceability, as pointed out in the following.

B. Bridging Law and Requirements

Existing requirements engineering frameworks generally rest on the idea of deriving the requirements for a software system from the analysis of the stakeholders' goals that the system-to-be will support once developed and deployed. This approach has demonstrated to be effective in successfully capturing strategic requirements. However, it hardly applies to the need of arguing about the compliance of such strategic requirements with the legal ones. As already pointed out, rights concern actions which intuitively are descriptions of the behavior wanted from the addressee actor, and can result in a goal or task of that actor. However, an action characterization in itself is not a goal neither a task for two main reasons. First, a goal is a state of the world *wanted* by an actor, whereas an action characterization is a state of the world *imposed* to the actor. Second, an action is a state of the world prescribed to an *abstract* actor —a *class* of actors, and as such it is also a *class* of actions. A goal, whereas is a specific state of the world wanted by a specific actor. This makes necessity to separate the concept of goal from the one of action characterization to avoid misleading shortcuts.

In Nòmos, to describe the concepts of the strategy, we adapt the *i** modeling framework [19]. Specifically, we use the *i** version as defined for the Tropos methodology [20]. Worth mentioning that this choice is arbitrary —other frameworks could be used or adapted to be used as well, as long as they provide primitives for modeling actors, goals, and relationships between actors. The *i** framework models a domain along two perspectives: the *strategic rationale* of the actors —i.e., a description of the intentional behavior of domain stakeholders in terms of their goals, tasks, preferences and quality aspects (represented as softgoals); and the *strategic dependencies*

among actors —i.e., the system-wide strategic model based on the relationship between the depender, which is the actor, in a given organizational setting, who "wants" something and the dependee, that is the actor who has the *ability* to do something that contributes to the achievement of the depender's original goals. An actor has the ability to achieve a goal if the actor has in the set of intentional elements that characterize it (such as sub-goals, tasks and resources) an element or a set of elements whose purpose is the achievement of the goal; or, if the actor can delegate the achievement of the goal to another actor — i.e., the dependee. The concept of ability is important because it allows to understand what are the characteristics of the specific actor existing in the domain, w.r.t the abstract actor addressed by the law.

With these ingredients, we are able to establish if a certain goal or task fits the characterization given by law, and to represent it in the model. In Figure 2, this is expressed with the concept of **realization** (class `Realization`), which puts in relation something that belongs to the law with something that belongs to the intentions of actors. This will be the starting point to argue about compliance of requirements models with law.

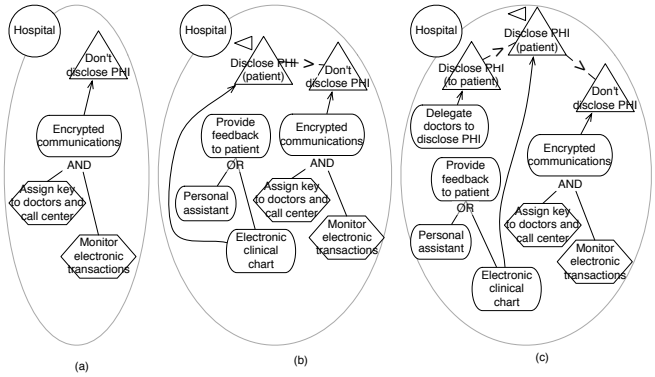


Fig. 3. An example: Law and Strategic Modeling using the Nòmos modeling language.

Figure 3(a) depicts excerpt model of a law fragment taken form the U.S. Health Insurance Portability and Accountability Act (HIPAA), and the goals' fulfilling law and tasks that operationalize such goals. The fragment contains the duty imposed to hospitals to keep patients' Personal Health Information (PHI) to be closed. The duty is fulfilled by the Hospital by setting up an encrypted electronic communication mechanism, which in turn is refined into the two leaf level tasks "Assign key to doctors and call center" and "Monitor electronic transaction". This choice means that the running system will need to support its processes with these activities. Alternatively (Figure 3(b)), the law gives hospitals the possibility to disclose patients' PHI, if the receivers are the patients themselves. The introduction of this new principle involves a possible change in the underlying processes supported by the system-to-be. Similarly, as in Figure 3(c), the introduction of the last principle —a duty, for the Hospital, to disclose such information to the patient upon request— further impacts on the supported processes, as it requires the hospital to receive the requests and to delegate somebody to disclose the data.

IV. Enhanced Law Modeling With VLPM

This section presents an approach where process models for procedures are modeled, and changes in laws are mapped in the models in order to highlight and review the impacts on processes and vice-versa. This allows for a stricter collaboration among the different stakeholders usually involved in BPR. These activities are mostly handled by the VLPM tool, as discussed below.

A. The VLPM Tool

VLPM [12] is a tool supported methodology for process modeling and re-engineering of PA, by providing a set of functions to synchronize models and XML representation of laws, thereby allowing traceability. The tool also supports an automatic generation of documentation in a human readable form (e.g., PDF or HTML), and of skeletons of law modifications based on the changes undergone by processes defined by the original law. The extended design of the framework makes the tool more flexible and functional in various areas. Among which we mention: support for different XML representation of laws, which are used by VLPM for linking process and laws; more flexibility in deployment — e.g., by allowing integration with freely available UML tools; integration with formal analysis techniques and methodologies for simulation and verification.

Figure 4 shows a high-level representation of the model elements, i.e. a meta-model for the VLPM tool. The diagram mainly shows the internal representation of the model elements. In the diagram, a process is realized as an observable activity executed by one or more actors. Actors can be extracted from the text of the law or can be defined manually. They are identified by means of an unambiguous identifier (extracted from the XML file containing the law information or manually specified) and a name. This could easily be extended in order to add more features. In the same way as actors, assets can be either extracted from the law or defined manually. If the assets are extracted from the law, we then store their initial states in the model and use our notation to define the changes that the assets undergo.

In addition to the modeling elements, we use a generic relationship elements to create specific sub-classes of relationship as shown in Figure 5. These are defined separately from the elements of the model. Actor-Actor relationships have different properties from Actor-Process relationships (e.g., the allowed stereotypes) and from Process-Process relationships. The association of a process with its executing actors is based on the static assignments of the responsibilities (set of roles, R) to the actors. This information can be extracted from the law or manually assigned after the actors have been identified. The use of an abstract relationship object allows us to create as many types of relationship as we need, with the only requirement of defining also a suitable translation of each

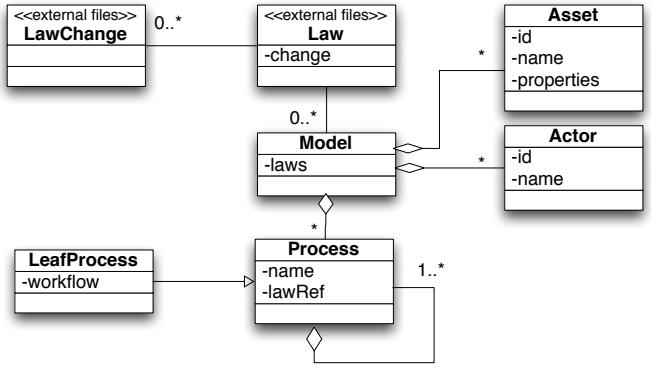


Fig. 4. The internal representation of our modeling elements.

relationship to UML. The model also explicitly support the Asset-Process relationships that define the semantics for the asset flows.

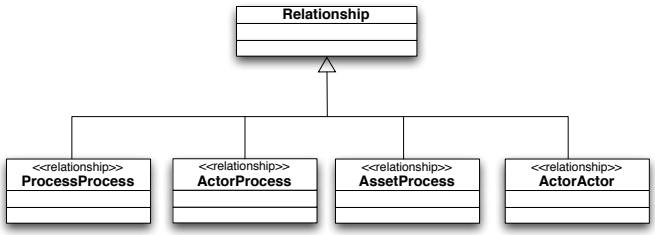


Fig. 5. Relationships among the modeling elements.

The model represents the static information of the business processes, while the dynamic properties (namely, asset transformation functions) are defined in a specific notation. The model is associated to the laws that regulate its business processes to allow the association of a single process with relevant law parts that define them. Notice, however, that the law is not included in the model. Although our model is designed to support XML format for laws representation, it can be easily extended to support other formats (see at the end of this section).

B. Methodology and Usage Scenario

We devised a methodology to automatically extract information from XML representation of laws and map them into process models. The core modeling elements are process, actor, asset, and relationships with triggering conditions. The methodology comprises of three steps.

The first step is the preparation of the data and structure of the model. Particularly, this step is responsible for identifying the actors, assets, stereotypes and terminologies, as well as responsible for collecting laws (the enumeration of laws which rule or influence the domain under analysis). The second step focuses on the use of UML *use case* diagrams to statically capture and analyze actors and processes independent from their execution. This is particularly important to breakdown processes hierarchically, to associate actors with responsibilities in the process breakdown structure, and to define and associate law paragraphs to processes in the use case diagram. Finally, the evolution of assets and processes are captured and analyzed using UML *activity diagrams*. Activity diagrams describe the processes workflow by emphasizing sequential and parallel activities (using the triggering conditions identified in step one) whose assets are needed and how their state evolve —i.e., how they change after being executing associated activities. The activity diagram also highlights the assets on which the processes operate. The connection between processes and assets are labelled with one or more of the following stereotypes: create, read, update, delete. In addition to the standard notation borrowed from the CRUD matrices [16], it is also possible to specify use, send and receive as stereotypes. This allows us to systematically translate into executable code (e.g., model checking) for further analysis, e.g., to perform procedural security analysis (such as, see in [21]).

The methodology also allows to link the laws and models. This particularly increases the traceability between laws and processes, with the goal of helping the law makers elaborate models in collaboration with software developers or process engineers, and understand the impact of law or process changes to their counterparts. This helps, first, to justify the existence of a particular process by providing a reference to the parts of the law that define it, which in turn allows us to link the process to all the constraints in the law that regulate it. Secondly, it allows to understand the impact of a change both in the law and in the process model. When a change is made to the law, on the one hand, being able to identify which processes are defined (or regulated) by the modified part of the law allows us to modify the process model accordingly. By looking at the model, it is then possible to determine what processes "interact" with the processes affected by the change in the law. The modification can then be propagated to all the relevant processes and makes the model up to date. On the other hand, the re-engineering of processes may result in a need to modify some parts of the law. Maintaining law-model traceability allows to automatically identify which parts of the law should be amended by tracing back to the parts of the law that originally defined the modified processes.

The points discussed above are supported by the VLPM tool, which has the following usage scenario (Figure 6 and see also in [12] for further detail):

1) A law written in natural language is marked with XML tags.
2) The user imports the law formatted in XML and VLPM generates a skeleton of the model. The user needs to verify and complete the generated model in order to have a reliable *as-is* view (i.e., a "process-tree" view) of the law. This model can be exported in various formats for documentation purposes (e.g., PDF).
3) The user imports an Explicit Text Amendment that modifies the law that has been previously modeled with VLPM. The tool highlights the impacts of the amendment on the law and on the model, allowing the user to focus on the affected parts of the model. This

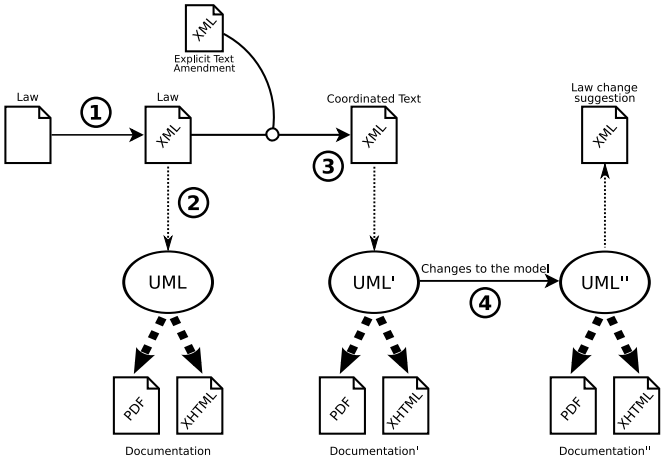


Fig. 6. Law modeling process handled by VLPM.

greatly simplifies the model revision process.

4) The user modifies the process model, re-engineering some processes. At this point documentation can be generated to be shared among the stakeholders and to compare the as-is and the to-be models. Moreover, VLPM can be used to generate the XML skeleton of a new law that amends the originally modeled law.

Furthermore, since the first version of VLPM focused on modeling the procedural aspects of legal documents that would allow to perform the necessary changes to the as-is business logic with the purpose. However, much information necessary for the reasoning cannot be easily extracted from the multitude of legal documents that regulate a domain without a background in jurisprudence. For this reason, the current version of VLPM (i.e., VLPM 2.0) exploits semantic markup on the laws to generate skeletons of business process models that can be traced back to the laws describing them. This allows lawyers and functional analysts to round-trip between laws and processes.

With VLPM 2.0, the components which produce the models from a set of legal documents provide some regulations for a certain domain. The information is layered hierarchically, where the bottom layer is the textual information and on top of which meta-data and structural information are added via the Akoma Ntoso [22], [23] markup XML format. We developed an OWL-DL ontology in order to add semantic information about processes described in legal texts, by extending the concepts of LKIF-core [24], [25] with a business process meta-model that borrows several entities from the BPMN meta-model [26]. We then defined some concepts that can effectively address our needs. We used Pellet Reasoner to identify and consolidate equivalences and other relations with LKIF-core classes. The VLPM 2.0 ontology is not a specification of the BPMN meta-model in OWL. Instead, it abstracts the core entities of a business process from the BPMN meta-model in order to obtain a smaller but more generic ontology. In the sense that a set of instances of the classes in such ontology can easily be transformed to BPMN as well as UML Activity Diagram (AD) [27]. Finally, we intend to support supplementary ontologies to allow the representation and modeling of other aspects of the domain, such as goal-oriented information.

V. Combining Nòmos and VLPM to Support BPR Scenario

Laws can express principles at different levels. Two levels are particularly apparent, high-level principles usually comprise of rules and requirements, and a set of procedural and/or operational level laws [28].

As we discussed previously, VLPM provides a robust environment to effectively manage the re-engineering of processes regulated by the set of operational laws. One significant limitation of the tool, however, is that it does not provide notations and means to represent the principles behind the procedures (or, better, motivating the procedures) and to reason about possible alternative implementation. Even from the business re-engineering point of view, such principles represent an essential part since they provide the framework and the constraints for the definition of new procedures and laws. This, in turn "moves" part of the re-engineering activity back to the "natural language" domain where inconsistencies and ambiguities might arise.

To overcome this problem, we propose the integration of goal-oriented methodology supported by Nòmos framework with the process modeling methodology supported by VLPM. The situation is depicted in Figure 7. On the left hand side, we have the law (possibly split in various documents) and typically describing general principles (e.g., "all citizens have the right to free health-care") and procedural and operational aspects (e.g., "to get free health-care you need a Social Security Number"). On the right-hand-side, we have to modeling techniques:

- Nòmos, in the upper part, provide a graphical notation and a methodology for modeling and reasoning about the high-level principles.
- VLPM, in the lower part, provide a graphical notation and methodology for modeling and reasoning about the procedures and lower-level operative principles.

Nòmos can represent the principles of the law via its constructs. In particular, as depicted in Figure 3(a) it is possible to represent the parts of a normative proposition such as the "Personal Health Information (PHI) closed", focusing on the actor, "Hospital", specified in the text of the law, also giving the possibility to specify the kind of right (in the case of the example a *duty*). Moreover, the framework allows to specify the goals that are induced by the text of the normative proposition, such as "set-up an encrypted electronic communication" and the specification of the actions that fulfill the goals and that represent the links to the procedural part of the law (in the example "Assign key to doctors and call center" and "Monitor electronic transaction"). Thanks to these representation, Nòmos maintains the complete knowledge about the principles at the bases of the operative part of the law, and of

the possible alternatives for the law fulfilling (see Figures 3(b) and 3(c)).

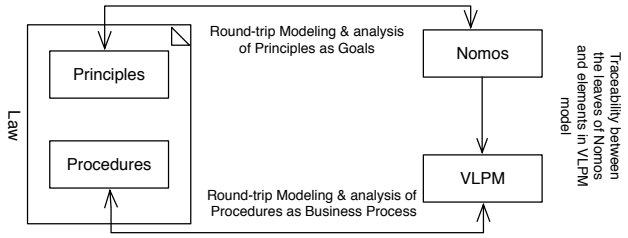


Fig. 7. The proposed approach for modeling and analysis of law.

VLPM extracts processes and actors from XML representation without the semantic knowledge that allows to reason on alternatives and here comes one of the essences of the Nòmos framework. Notice that the leaves of the Nòmos model can be analyzed on their fulfillment and on their compliance with the actual norm/law. There are also correlations between leaves in Nòmos and activities in VLPM. Thus, it is straightforward to say that such leaves can help enriching the VLPM "process-tree" —hierarchical decomposition of processes and actors using UML use cases— with more semantics on the management of activities in the VLPM model.

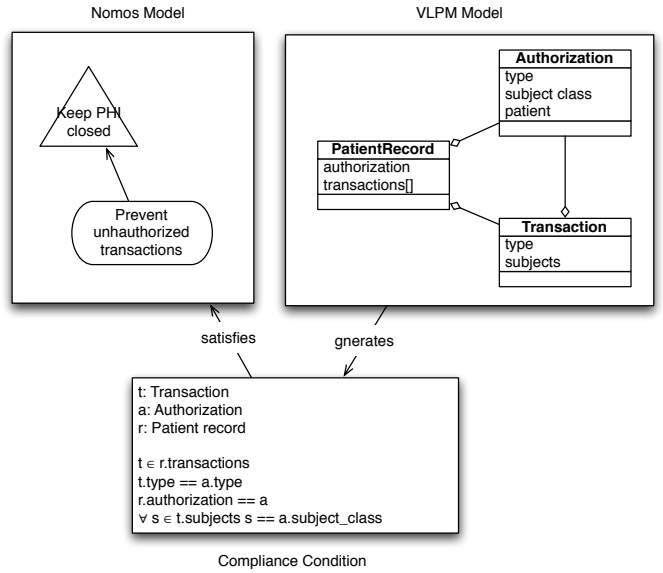


Fig. 8. A simple illustration using the proposed approach.

Figure 8 illustrates how the approach actually works. The Nòmos model contains the description of the principles that should be respected by the stakeholders. Moreover, it contains the description of the strategic goals that the stakeholders develop to be compliant. The presence of goals allows to make a domain-dependent analysis, which takes into account their specific objectives and needs, besides those of the law. The VLPM model contains a description of the system architecture. More precisely, it contains a description of that part of the system whose definition has been extracted from the annotated law. In the bottom part of the figure, the compliance condition is depicted. Basically, the compliance condition consists of a representation of the achievement condition of the goals of the Nòmos model. The condition says that the compliance goal "Prevent unauthorized transactions" is satisfied when the system is in a state represented by the value of the three variables t, a, and r, where t is an instance of the class `Transaction`, a is an instance of the class `Authorization`, and r is an instance of the class `PatientRecord`. The state represented by the condition is such that the transaction of certain data is linked with the patient's authorization to transmit that data. In the picture, this is considered true if:

- the transaction has also been associated to an authorization ($r.authorization == a$);
- the authorization is specific for that type of transaction ($t.type == a.type$);
- the transaction has actually been recorded into at least one patient record ($t \in r.transactions$) – i.e., no transaction happen, which are not registered and do not respect the other conditions;
- the subjects, among which the data is exchanged match the role type declared in the authorization for that transaction ($\forall s \in t.subjectss == a.subject\_class$).

Given this compliance condition between any possible state of the system, described in terms of the values of its variables and the principles expressed by the law, it is possible to exhaustively check for states that are allowed by the system but not acceptable for the law.

An important aspect to highlight is the traceability offered by the two approaches are complimentary. For example, if you decide to remove a process from the UML model that corresponds to one of the leaves of the Nòmos model, the Nòmos framework can trace up to the root goal and check if this action is complaint with the actual norm (from which the leave is derived). It is worth mentioning that when a new process is added to the model, VLPM generates a list of suggestions that can be used to produce an explicit text amendment from the changes undergone by the model, thus allowing the law to be realigned to the model. This can be further refined and enhanced by using the power of Nòmos analysis.

This conforms the connection between the Nòmos model and the VLPM model. The leaves of the Nòmos model, in fact, correspond to the procedures of the VLPM model. This provides a framework that allows to trace and reason either top-down, from the principles to the implementation, or, vice-versa, bottom-up, from a change in the procedure to the principles.

VI. Related Work

Several strategies have been proposed in the literature to understand, model, and analyze business process models. Three aspects are central in these approaches. The first is tools used for creating (business) process models. Second, notations used to represent the modeling elements and concepts. Third, techniques used for formally specifying and verifying how such models respect the intended goals.

With respect to modeling of business processes, for instance, various works in the past have been proposed for modeling business processes. These approaches span from workflow nets to event-based process chains, from flow-charts diagrams to UML Activity diagrams (ADs) and Business Process Execution Languages for Web Service (BPEL4WS)[29] and several other works such as [30], [31], [32]. In particular, [33], [34], [35] widely discussed the usage of UML ADs for modeling business processes as well as workflow modeling and specifications.

While these works demonstrated their usage scenarios for the modeling, specifying, and analyzing business processes and workflows of complex system, the attempt to model laws and procedures, as well as to perform formal analysis in favor of the public administration (PA) is not satisfactory.

In recent years, however, a number of governments have been adopted such techniques to support their PAs. Works like [3], [2], [36], [37] particularly have been discussed BPR support for public healthcare services by identifying different levels of process support and by distinguishing among generic process patterns. The use of BPR for better government has also been discussed earlier by the U.S. federal government and the U.S. Department of Defense and its use in taxation in [38]. The importance of modeling in the legal framework and documenting the knowledge about the legal constraints within the process model itself is stated in [4].

In particular, Olbrich and Simon in [39] discussed an approach based on event-driven process chains and suggested how to translate law paragraphs into process models using the Semantic Process Language (SPL). Their main goal is to the visualization and formal modeling of a legally regulated process. The interesting aspect of this work is not only the consideration of the given law when developing business process models, but also the explicit derivation of a process structure which is implicitly specified within the paragraphs themselves using the SPL. The SPL enabled them to articulate language structures into executable workflow models, using Petri Nets. The presented approach could provide means for verifying whether process-like behavior fulfills the selected paragraphs formally. Related to processes and their verification, in [40] the authors propose a UML-based approach to define, verify, and validate organizational processes, especially in the context of software process improvement and the CMMI (Capability Maturity Model Integration) framework.

Related to goal-oriented approaches for modeling and representation of laws and with the compliancy of set of requirements to laws. Three of them are particularly relevant for our work. In [8], the authors used KAOS as a modeling language for representing objectives extracted from regulation texts. Such an approach is based on the analogy between regulation documents and requirements documents. In [9], Goal-oriented Requirement Language (GRL) to model goals and actions prescribed by laws and exploit Use Case Maps (UCM) to describe the impact of laws on the business processes is discussed. This work is founded on the premise that the same modelling framework can be used for both regulations and requirements. In [28] is shown that two levels exist in legal systems: the Rule level, which gives prescriptions in an Event-Condition-Action (ECA) style; and the Requirements level, which expresses desirable state of affairs to be achieved by addressees. It also argues that the requirements level cannot exist alone: it depends on the rule level for actuation and enforcement. However, it tackles only with the requirements level while discussing the integration of laws into enterprise configurations.

Breaux et al. in [41] develop a systematic process called semantic parameterisation using the Cerno framework [42]. The approach consists of identifying in legal text restricted natural language statements (RNLSs) and then expressing them as semantic models of rights and obligations [10] (along with auxiliary concepts such as actors and constraints). Secure Tropos [43] is a framework for security-related goal-oriented requirements modeling that, in order to ensure access control, uses strategic dependencies refined with concepts such as: trust, delegation of a permission to fulfill a goal, execute a task or access a resource, as well as ownership of goals or other intentional elements. The main point of departure from Nòmos is that the Nòmos use a richer ontology for modeling legal concepts, adopted from the literature on Law. Additionally, the legal models one builds using Nòmos is different from the mentioned usage —i.e., Nòmos allows to check compliance between an i* model of system requirements and a model of a law fragment.

VII. Conclusion

The application of BPR and goal-oriented in law modeling and analysis can facilitate the work of PAs by favoring the involvement of citizens in (the law) decision process. The definition of strict constraints for the structure of a law facilitates its readability and editing, but —in the case of laws definition procedures —the use of (visual) representations, their modeling, and formal reasoning can take this even further.

This paper proposed an approach intended to provide systematic support for modeling and analyzing laws. Our approach combines two existing complementary frameworks that tackle the discussed concerns in different levels of abstractions. While one (i.e., Nòmos) exploits goal-oriented approach, the other (i.e., VLPM) focuses on the use of UML-based BPR approach. We emphasized on the integration of these two approaches for realizing principles, procedures, as well as operational aspects of the law, and for developing a system that can maintain and support the laws.

The resulting analysis method of the Nòmos approach takes advantage of two key ideas, namely the concept of intentional compliance to verify law-compliance of requirements models, and the idea of combining a law model with an intentional model of requirements for preserving the explicit representation of compliant alternatives resulting from goal analysis. In contrast, VLPM is based on well established technologies for legal knowledge representation such as LKIF and Akoma Ntoso, including process and goal-oriented ontologies. The current VLPM framework will have the possibility to play

an important role in the "ICT for law" initiatives and eventually become an actual tool for the improvement of Public Administrations.

In summary, the proposed approach has given promising clues to trace and reason laws either from the principles to the implementation, or, vice-versa, from a change in the procedure to the principles. Although, the snippet example used for our proof-of-concept is not complete, we believe that the proposed approach can be refined and (re-)used to model and analyze different laws, as long as the laws describe both the principles and procedures. Namely, as long as the the laws define, regulate or in some way affect procedures, e.g., PA procedures, company policies that need to comply with certain regulations. Moreover, the implementation of the approach in terms of a tool is not discussed. However, we are currently working on enriching the two frameworks because this would allow us to develop a machinery for the combined activities. We are also looking for a real case study to evaluate our approach.

REFERENCES

[1] A. Villafiorita, K. Weldemariam, A. Susi, and A. Siena, "Modeling and Analysis of Laws Using BPR and Goal-Oriented Framework," *International Conference on the Digital Society*, pp. 353–358, 2010.

[2] W. C. Leslie Willcocks and S. Jackson, "Radical Re-Engineering and Information Systems: Evidence from UK Public Services," in *Fifth European Conference in Information Systems*. Cork, 1997.

[3] V. B. Marcel Thaens and H. van Duivenboden, "Business Process Redesign and Public Administration: a Perfect Match?" in *Taylor, J.A., Snellen, I.Th.M. and Zuurmond, A. (Eds.): Beyond BPR in Public Administration: An Institutional Transformation in an Information Age*. IOS Press, Amsterdam, 1997, pp. 15–36.

[4] P. Alpar and S. Olbrich, "Legal Requirements and Modelling of Processes in e-Government," *Electronic Journal of e-Government*, vol. 3, 2005.

[5] A. Siena, "Engineering Law-Compliant Requirements. The Nòmos Framework." Ph.D. dissertation, Department of Information Engineering and Computer Science, University of Trento, Trento, March 2010. [Online]. Available: http://eprints-phd.biblio.unitn.it/230/

[6] R. Kowalski and M. Sergot, "Computer Representation of the Law," in *Proceedings of the 9th international joint conference on Artificial intelligence - Volume 2*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1985, pp. 1269–1270.

[7] T. J. M. Bench-Capon, G. O. Robinson, T. W. Routen, and M. J. Sergot, "Logic Programming for Large Scale Applications in Law: A Formalisation of Supplementary Benefit Legislation," in *Proceedings of the 1st international conference on Artificial intelligence and law*, ser. ICAIL '87. New York, NY, USA: ACM, 1987, pp. 190–198.

[8] R. Darimont and M. Lemoine, "Goal-oriented Analysis of Regulations," in *ReMo2V*. CEUR-WS.org, 2006.

[9] S. Ghanavati, D. Amyot, and L. Peyton, "Towards a Framework for Tracking Legal Compliance in Healthcare," in *Proceedings of the 19th international conference on Advanced information systems engineering*, ser. CAiSE'07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 218–232.

[10] T. D. Breaux, A. I. Antón, and J. Doyle, "Semantic Parameterization: A Process for Modeling Domain Descriptions," *ACM Trans. Softw. Eng. Methodol.*, vol. 18, no. 2, pp. 1–27, 2008.

[11] A. Siena, J. Mylopoulos, A. Perini, and A. Susi, "A Meta-Model for Modeling Law-Compliant Requirements," in *Proceedings of RELAW workshop at RE 2009*. IEEE Computer Society, 2009.

[12] A. Ciaghi, A. Mattioli, and A. Villafiorita, "VLPM: A Tool to Support BPR in Public Administration." in *ICDS*. IEEE Computer Society, 2009, pp. 289–293.

[13] A. Ciaghi, A. Villafiorita, and A. Mattioli, "A Tool Supported Methodology for BPR in Public Administrations," *International Journal of Electronic Governance*, vol. 3, no. 2, 2010.

[14] R. Hoekstra, J. Breuker, M. Di Bello, and A. Boer, "LKIF Core: Principled Ontology Development for the Legal Domain," in *Proceeding of the 2009 conference on Law, Ontologies and the Semantic Web*. Amsterdam, The Netherlands: IOS Press, 2009, pp. 21–52.

[15] L. Sommerville, *Software engineering (5th ed.)*. Addison Wesley Longman Publishing Co., Inc., 1995.

[16] A. Mattioli, "Analysis of Processes in the Context of Electronic Election," Master's thesis, University of Trento, Italy, 2006, in Italian.

[17] L. Wenar, "Rights," in *The Stanford Encyclopedia of Philosophy*, fall 2010 ed., E. N. Zalta, Ed., 2010.

[18] W. N. Hohfeld, "Some Fundamental Legal Conceptions as Applied in Judicial Reasoning," *Yale Law Journal*, vol. 23, no. 1, 1913.

[19] E. S.-K. Yu, "Modelling Strategic Relationships for Process Reengineering," Ph.D. dissertation, University of Toronto, Toronto, Ontario, Canada, 1996.

[20] A. Susi, A. Perini, J. Mylopoulos, and P. Giorgini, "The Tropos Metamodel and Its Use." *INFORMATICA*, vol. 29, no. 4, pp. 401–408, 2005.

[21] K. Weldemariam and A. Villafiorita, "Formal Procedural Security Modeling and Analysis," in *International Conference on Risks and Security of Internet and Systems*, ser. CRiSIS '08. Washington, DC, USA: IEEE, October 2008, pp. 249–254.

[22] United Nations Department of Economic and Social Affairs. (2010) Akoma Ntoso Framework: Architecture for Knowledge-Oriented Management Of African Normative Texts Using Open Standards and Ontologies. [Online]. Available: http://www.akomantoso.org/

[23] F. Vitali and F. Zeni, "Towards a Country-Independent Data format: the Akoma Ntoso Experience." in *Proceeding of the V Legislative XML Workshop*, 2007, pp. 239–252.

[24] ESTRELLA project. (2010) LKIF-Core Ontology: A Core Ontology of Basic Legal Concepts. [Online]. Available: http://www.estrellaproject.org/lkif-core/

[25] R. Hoekstra, J. Breuker, M. D. Bello, and A. Boer, "The LKIF Core Ontology of Basic Legal Concepts," in *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques (LOAIT 2007)*, P. Casanovas, M. A. Biasiotti, E. Francesconi, and M. T. Sagri, Eds., June 2007.

[26] "Business Process Modeling Notation (BPMN) Version 1.2," January 2009. [Online]. Available: http://www.omg.org/spec/BPMN/1.2/PDF

[27] G. Booch, J. Rumbaugh, and I. Jacobson, *Unified Modeling Language User Guide, The (2nd Edition) (Addison-Wesley Object Technology Series)*. Addison-Wesley Professional, 2005.

[28] W. Hassan and L. Logrippo, "Requirements and Compliance in Legal Systems: A Logic Approach," *Requirements Engineering and Law*, pp. 40–44, 2008.

[29] M. B. Juric, *Business Process Execution Language for Web Services BPEL and BPEL4WS 2nd Edition*. Packt Publishing, 2006.

[30] J. Lee, "Goal-Based Process Analysis: A Method for Systematic Process Redesign," in *Proceedings of the Conference on Organizational Computing Systems*, ser. COOCS '93. New York, NY, USA: ACM, 1993, pp. 196–201.

[31] M. M. Lehman, "Process Modeling —Where Next," in *Proceedings of the 19th Internaional conference on Software Engineering*, ser. ICSE '97. New York, NY, USA: ACM, 1997, pp. 549–552.

[32] V. Hlupic, "Business Process Modelling Using Discrete Event Simulation: Potential Benefits And Obstacles For Wider Use," *International Journal of Simulation: Systems, Science and Technology*, vol. 7, pp. 62–67, 2003.

[33] M. Dumas and A. H. M. t. Hofstede, "UML Activity Diagrams as a Workflow Specification Language," in *Proceedings of the 4th International Conference on The Unified Modeling Language, Modeling Languages, Concepts, and Tools*. London, UK: Springer-Verlag, 2001, pp. 76–90.

[34] N. Castela, J. M. Tribolet, A. Silva, and A. Guerra, "Business Process Modeling with UML," in *ICEIS (2)*, 2001, pp. 679–685.

[35] R. Eshuis, "Symbolic Model Checking of UML Activity Diagrams," *ACM Trans. Softw. Eng. Methodol.*, vol. 15, no. 1, pp. 1–38, 2006.

[36] G. Greco, A. Guzzo, and L. Pontieri, "Mining Hierarchies of Models: From Abstract Views to Concrete Specifications," in *Business Process Management*, 2005, pp. 32–47.

[37] R. Lenz and M. Reichert, "IT Support for Healthcare Processes," in *Business Process Management*, 2005, pp. 354–363.

[38] U.S.A. National Performance Review, "Executive Summary — Creating a Government that Works Better and Costs Less," 1993. [Online]. Available: http://govinfo.library.unt.edu/npr/library/nprrpt/annrpt/redtpe93

[39] S. Olbrich and C. Simon, "Process Modelling towards e-Government – Visualisation and Semantic Modelling of Legal Regulations as Executable Process Sets," *Electronic Journal of e-Government*, vol. 6, 2008.

[40] N.-L. Hsueha, W.-H. Shen, Z.-W. Yanga, and D.-L. Yanga, "Applying UML and software simulation for process definition, verification, and validation," *Information and Software Technology*, vol. 50, pp. 897–911, 2008.

[41] T. D. Breaux, M. W. Vail, and A. I. Anton, "Towards Regulatory Compliance: Extracting Rights and Obligations to Align Requirements with Regulations," in *RE'06: Proceedings of the 14th IEEE International Requirements Engineering Conference*. Washington, DC, USA: IEEE Computer Society, 2006, pp. 46–55.

[42] N. Kiyavitskaya, N. Zeni, J. R. Cordy, L. Mich, and J. Mylopoulos, "Cerno: Light-Weight Tool Support for Semantic Annotation of Textual Documents," *Data Knowl. Eng.*, vol. 68, no. 12, pp. 1470–1492, 2009.

[43] P. Giorgini, F. Massacci, J. Mylopoulos, and N. Zannone, "Requirements Engineering meets Trust Management: Model, Methodology, and Reasoning," in *ITRUST-04*, ser. LNCS, vol. 2995. SVG, 2004, pp. 176–190.

Integrating Future High End Computing and Information Systems Using a Collaboration Framework Respecting Implementation, Legal Issues, and Security

Claus-Peter Rückemann
Leibniz Universität Hannover, Hannover, Germany
Westfälische Wilhelms-Universität Münster (WWU), Münster, Germany
North-German Supercomputing Alliance (HLRN), Germany
Email: ruckema@uni-muenster.de

***Abstract*—This paper gives an extended overview of challenges creating complex integrated information and computing systems. It covers implementation, legal, and security issues with these processes and how the overall complexity can be reduced using collaboration frameworks with Distributed and High Performance Computing resources in natural sciences disciplines for building integrated public / commercial information system components within the e-Society. Focus is on using a collaboration framework for implementing computing resources, interfaces for data and application interchange, based on current developments regarding infonomics systems, last years case studies within the long-term GEXI project, and Active Source. A suitable framework base has been created over the last years, being used for a number of scenarios in research environments, using High End Computing resources. Application of these methods for commercial service structures affords the consideration of various legal, security, and trust aspects. In collaboration with international partners from natural sciences, industry, economy, and education the framework has been found the solution to overcome the legal cooperation barrier. Established on this work, international cooperations are currently built. In addition, this paper presents two major implementation case studies in order to show the application of the collaboration framework, one for environmental and energy exploration information, computing, and resources management and another for epidemology information systems.**

Keywords–Legal Frameworks; Collaboration Management; Implementation; Legal Issues; Security; Distributed Systems; High Performance Computing; Grid-GIS house; e-Science; Geoscientific Information Systems.

I. Introduction

Todays information system design, development, implementation, and usage are in many cases characterised by dynamism and fast varying means, heterogeneous content, access and information security in complex environments, short-term financing, and individual architectures. About over a decade now, the amount of information available as well as the computing power has been continuously increasing, but there is no integrated information-computing system actually really bringing these vast resources together.

Research on overcoming these shortcomings for international collaboration management is going on for the last years [1]. Over the last years a long-term project, Geo Exploration and Information (GEXI) [2] for analysing case studies, has examined chances to overcome the deficits.

This paper presents the current results with a collaboration framework that has been developed and successfully used as a solution for various cases. It delivers the results collected from a study taken on participating national and international collaboration projects regarding the sections High Performance Computing (HPC), Distributed Computing (DC) and services, and natural sciences. There is a number of factors limiting the vigors that are devoted on the development of integrated systems. These are the problems with distribution of valuable resources like High End Computing (HEC) needed to be integrated on one hand and the legal diversities as with automation, personalisation, security, and differences in professional, national, and international context on the other. As resulting from the developments of a suitable framework for DC and HPC, this work provides the legal complement for the technical and scientific base currently used for various international collaborations, building information, processing and decision making systems. Two case studies are presented in order to discuss the different aspects of “trust in computing” and “trust in information” emphasis.

II. Motivation

Geoscientific information systems belong to the most advanced information systems available today. A driving force behind the development besides public interests is from applied natural sciences, exploration and energy, resources, and environmental management. Oil and gas, climatology, aerospace and automotive industry for example depend on privacy for their computations. Neither services nor security in general will change the advised behaviour within the next decade. It is uninviting to expect the driving force from the resources engineering approach only. It is a common misbelief that industry will use any distant foreign HEC resources for economically interesting or critical calculations, for computing or storage in context with their strategic data or with production. The research projects of the last years have shown that we need a legally integrated comprehensive

collaboration framework for complex modular design and development [3] as well as new methods and algorithms on the geoinformation side [4] regarding distributed resources and secure communication.

This paper is organised as follows. Section three presents preliminary work and cases studies. Sections four and five describe legal frameworks, regulations and name the problems addressed. Section six describes the new collaboration framework. Section seven explains the status of the implementation regarding the participating key player topics. Section eight reports evaluation and consequences for primary topics on technology and legal issues and the collaboration portal. Section nine gives an extended presentation of selected case studies, implemented following the collaboration framework. The focus is on implementation, legal issues, and security aspects faced by the resources, services, and disciplines columns. Section ten shows the lessons learned and Section eleven summarises the conclusion and outlook on future work.

III. Preliminary work and case studies

This discussion of legal issues is the result of the successful work of the last years, in the disciplines of geosciences application architecture [5] (Active Source), preparing an interdisciplinary Grid-GIS house framework [6], project case studies, and the adjunctive configuration of various HEC (HPC, Distributed, Grid, Cloud, and GPU Computing) resources over the last three years [4]. Work is ongoing for developments in cooperation with international industry and economy partners [7], [3].

The analysis of these case studies showed that the coordinated cooperations have a strong need to address the legal and security base for handling critical data (e.g. business relevant development and exploration data), computing and geo-processing as well as components used.

At todays level of information integration, with the overall complexity of information and decision making systems, there is a neccessity for building flexible and extendable information systems for cognostic purposes that consider aspects of security and economy in complex environments. For future cooperations and implementations it is most important to focus on legal issues regarding the frameworks.

IV. Legal frameworks and regulations

As in this context, there exists no collaboration development framework, there is a number of partially interesting international and national legal topic-frameworks (Table I) for information systems, regarding content and structure.

Table I
Legal frameworks with geosciences.

Name	*Framework and Reference*	
GMES	Global Monitoring for the Environment and Security	[8]
GEOSS	Global Earth Observation System of Systems	[9]
	of the GEO (Group on Earth Observations)	[10]
SEIS	Shared Environmental Information System	[11]
GSDI	Global Spatial Data Infrastructure	[12]
INSPIRE	Infrastructure for Spatial Information in Europe	
	directive (2007/2/EC)	[13]
GDI-DE	Geodateninfrastruktur Deutschland	[14]
PSI/EPSI	Public Sector Information directive /	
	European Public Sector Information	[15]

Besides these frameworks there is a number of laws and legal regulations regarding geo data in Germany: copyright law (Urheberrechtsgesetz, UrhG), data security and privacy law (Bundesdatenschutzgesetz, BDSG), freedom of information law (Informationsfreiheitsgesetz, IFG), law on the reuse of information from public institutions (Gesetz über die Weiterverwendung von Informationen öffentlicher Stellen, Informationsweiterverwendungsgesetz, IWG), environmental information law (Umweltinformationsgesetz, UIG), law on accessing digital geo data (Gesetz über den Zugang zu digitalen Geodaten, Geodatenzugangsgesetz, GeoZG). These regulations do concern many aspects, contributors and participants of a modern information system, content as well as implementation, access, and usage.

V. Problems addressed

Currently there exist traditional information structures, old non-intelligent applications and data formats, but the data-barrier at various media is still omnipresent. Although the earliest of the fundaments of the named academic frameworks are dating back into the year 1998 there is still no information system available integrating the technical and legal diversities. But all of these frameworks can be seen complementary to comprehensive information system frameworks for geo-computing and geo-processing. What we target to, is legally conform modular applications and data formats for integrability of all resources.

A complementary effort is necessary for integrating industry, economy, and legal expertise into this process of creating a next generation information system framework at an international level. Further industrial and economic impulses originate from exploration, geosciences, energy-sciences, climatology, and education for handling data on resources management, observation, environment, biodiversity, weather, medicine.

In these disciplines many applications do need an integration of information systems with simulation and virtual reality. Not only that oil and gas industry, insurances, town planning, tourism industry and many others do have strong needs to be integrated into this processes. The integration

will provide chances for new insights into complexity of the environmental systems. Therefore, primary goals are implementations based on the collaboration framework:

- integration of academia, industry, economy, law,
- modularisation of system components,
- structuring of data and information,
- integration of HEC and storage resources,
- legally conform georeferencing of data and objects,
- licensing (e.g. topography, remote sensing),
- personalisation of information and services.

VI. Collaboration Framework

Illustrating the directions of integrating and co-developing large collaboration target frameworks and applications for service-oriented DC and HPC, Figure 1 shows the columns of the infonomics system and Figure 2 shows the dependencies of market and services (green colour, shingle and cross pattern), computing services (red colour, brick pattern), HPC and distributed resources (blue colour, gravelly pattern), and resources to be provisioned or developed (gold colour). The proposed Computing Industry Alliance ([4], Leadership in Research consortium) will be a suitable umbrella organisation for distributed and HPC and geo-exploration sciences. The framework described is an example currently building the base for creating efficient interdisciplinary industry research cooperations for implementing the next generation of dynamical applications on distributed and HPC resources based on the "Grid-GIS house" [6]. Interests to force this development exist, not only in the Gulf of México region but as well in Russia and Saudi Arabia.

Resulting from the GEXI project [2] started in the year 1996 as a public and private support network, the components and mechanisms have been topic of several information science, HPC, Grid Computing cooperations and European activities of the last years [16].

Three key player collaboration sections from HPC and Distributed and Grid Computing, from services and technical development, and from geosciences and exploration are currently building the next generation of information and computation system.

With support of international partners from geosciences, high end computing industry, economy, and education an extended Grid-GIS house [6] for the geosciences and exploration disciplines has been created this year [3] and legal aspects are currently discussed for further cooperation.

VII. Status of the Implementation

A. Infonomics system and interactions

The entirety of the essential columns of the framework, geosciences and energy-sciences, Distributed Computing and services, and HPC forms a well balanced infonomics system. The necessary interactions for the information and computing systems build the interfaces for the columns of the infonomics system (Figure 1). For integration into the

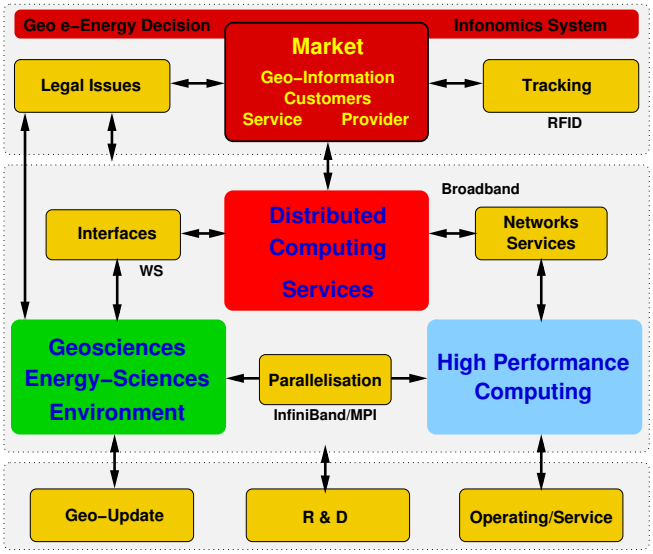


Figure 1. Columns of the infonomics system.

market, suitable services (e.g. Software as a Service, SaaS) will be provisioned as geosciences and e-Energy seek a complex information and decision making environment. Efficiency and effectivity is important from various perspectives: Fast networks (broadband, InfiniBand) are needed for distributed resources and HPC, interfaces for services and scientific applications, complementary to the simple OGC Web Processing Service (WPS), as well as parallelisation for geoscience algorithms is currently expedited with the collaboration framework in order to be employed on HPC resources.

Supplies for the natural sciences disciplines will be done with updates of data and algorithms and the computing resources do need a continuous operating as research and development are essential for interaction between the columns of the system. As the market does not only want to "trade" electronic goods, a suitable coupling with the information systems is necessary for physical identity tracking and monitoring, e.g., RFID for container cargo and Intelligent Transport Systems (ITS). Legal issues regarding all of these topics are omnipresent, for the columns and for the market.

B. Legal focus points

The entirety of these aspects describes a next generation "Information and Decision Making Environment" for the future internet, containing electronic and physical operations like information system components and logistics and tracking support for objects and goods.

Figure 3 shows the important legal focus points and dependency relations.

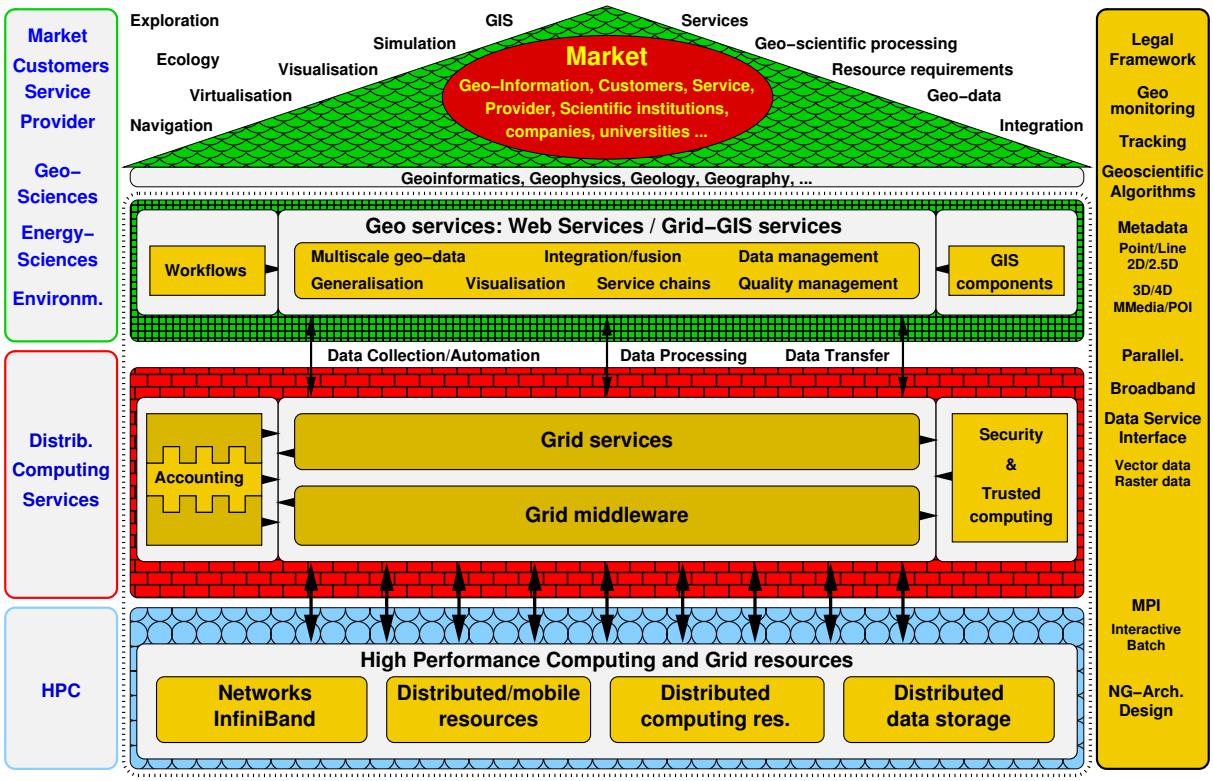


Figure 2. Collaboration framework for geosciences / exploration and HEC ("Grid-GIS house").

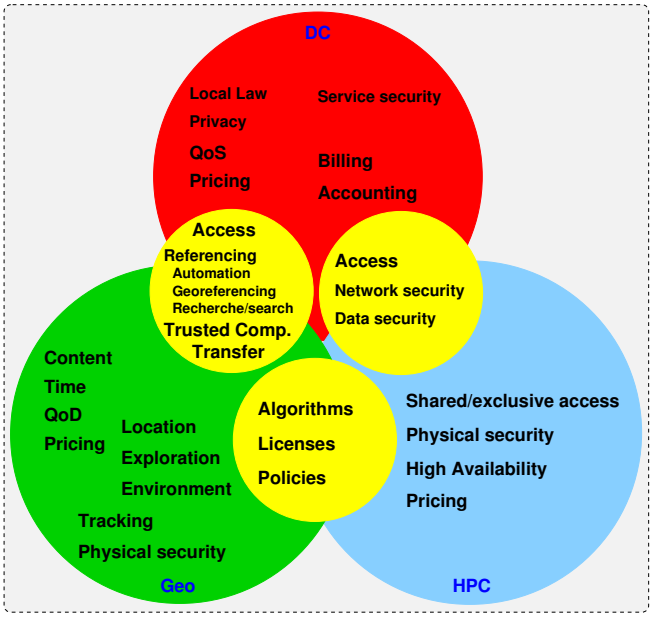


Figure 3. Legal focus points with GEXI.

As some related issues (yellow colour) strongly overlap with different columns they should be worked on in a collaboration development between disciplines, providers, and computing industry.

There have been implementations regarding a number of issues within case studies and future topics have been identified. Table II shows a compact overview of the implementation focus points and the current directions (✓ mostly done, ➠ future tasks within project). As the table points out, a number of issues is already done others will have to be worked on, for integrating new features needed for the next stage for an upheaval in system complexity. Some other issues will have to remain undone work with regard to standardisation as they afford an individual configuration, parametrisation, and optimisation, like access to HPC resources or because there are suitable solutions like MPI for parallelisation.

The number of legally safe application scenarios in focus, working with these resources and technologies should be increased by the participating academia and industry. These scenarios include commonly shared interactive distributed resources usage for computation and information storage and processing, distributed among international partners with different legal conception and regulations regarding privacy and anonymity.

There is a number of selected topics, being in the focus of current implementation and usability studies. The following passages show some topics important for the interdisciplinary context of the case studies.

C. Security

Physical and logical aspects and features, hazards and threats have been considered within the last years.

Table II
IMPLEMENTATION FOCUS POINTS AND DIRECTIONS.

Issue	*Implementation & Status*		*Future*
Access	individual	✓	➠ standardisation
	Distributed Comp.	✓	(—)
	High Perf. Comp.	✓	(—)
Trusted Comp.	sandbox, security policies	✓	➠ porting, PKI
Transfer	encryption / keys pairs	✓	➠ WebServices
	broadband	—	➠ industry/federal
QoS	categorisation	—	interdisciplinary
Content	composite data	—	interdisciplinary
	vector, raster	✓	➠ standardisation
	attribute data	✓	➠ standardisation
	event data	✓	➠ standardisation
	references	✓	➠ standardisation
Licensing	individual	—	➠ license mgmt.
QoD	level cognostics	—	interdisciplinary
Referencing	automation	—	➠ scripting
	georeferencing	—	➠ categories
	recherche / search	—	➠ spec. engines
Algorithms	parallelisation	✓	➠ standardisation
	loosely coupled	✓	(—)
	MPI, OpenMP, Java	✓	(—)
Tracking	phys. identification	—	➠ RFID
Accounting	non-commercial acc.	✓	➠ undisputable
	integrated solution, SGAS	✓	➠ modules
Pricing	individual, flatrate	✓	➠ compound units
Billing	individual, flatrate	✓	➠ cumulative

- Restrictions on hardware usage.
- Restrictions on access.
- Trust in information.
- Trust in computing.
- Trusted scripting.
- Inter Process Communication (IPC).
- Active Source methods.
- Accounting security, bidirectional and undisputable.
- Sandboxing and policies.

In the implementation case studies these aspects had to be handled in order to implement real-life systems. This will be discussed in some of the next sections.

D. Sandboxing policies

Various sandboxing mechanisms have been experimented with. The Tcl architecture supports very flexible features for sandboxing and policy implementation. With dynamical client-server applications the Tcl plugin supports Tcl/Tk applets, so called Tclets. The Tcl plugin implements the standard Safe-Tcl subset and defining new policies. For the Safe-Tcl interpreter various commands can be removed from the Tcl interpreter by configuration, used to run Tcl applets. A limited version of Tk has been added. This sandboxing can be used in the most flexible way with high level languages and other scripting languages. With the case studies, control for the following commands has been found most important, in order to gain trustable modules.

- exec (execute programs),
- load (dynamically load shared libraries implementing C or Tcl language commands),
- open (open a file, restricted open-read-only version available),
- send (send Tcl commands to other applications),
- cd (change directory),
- socket (open a network socket),
- source (load script files),
- exit (terminate a process).

Tk images cannot be created or read from files. Image create photo commands take strings of base64 encoded images instead. Further commands are handled with the Safe-Tcl.

- wm (window manager control),
- toplevel (create toplevel windows),
- menu (display a menu),
- tk (set and query Tk application names),
- tkwait (block on events),
- bell (ring terminal bell),
- clipboard (access the clipboard selection),
- glob (match file names in a directory),
- grab (grab the cursor),
- pwd (query present working directory).

These and additional functions and commands can be configured for the executing sandbox environments.

E. Accounting implementation

An integrated accounting and billing approach has been developed in the last years. The SweGrid Accounting System (SGAS) [17] has been considered most useful for Distributed and High Performance Computing. It supports scalable resources and capacity allocation [18], [19], decentralised concepts for fairshare scheduling [20], OGSA-based bank service [21], distributed usage logging [22], and support for federated cloud infrastructures [23]. The integrated accounting and billing approach supports market-ready secure, transparent, and flexible resources management.

The following section evaluates the current status and describes the lessons learned for implementation, technical and legal consequences.

VIII. EVALUATION AND LEGAL CONSEQUENCES

Some aspects like loosely coupled (with Grid, Cluster, and HPC) as well as MPI parallelisation for Massively Parallel Processing (MPP) and Symmetric Multi-Processing (SMP) have been successfully implemented and used for various purposes (e.g. in the projects Condor-network, ZIVGrid, ZIVHPC, ZIVSMP, HLRN) [6], [4]. It has been proven viable to use a collaborative implementation strategy to integrate individual solutions for making long-term investments sustainable. International work is currently done for parallelising application-triggered algorithms for interactive

use on huge HPC resources. Focus is on management, exploration, and environmental applications for geosciences, energy management, and information sciences. Invited industry partners are currently implementing parallelised application suites. From the legal point of view, protocols and exclusive commercial use of resources have to be implemented for reliable use of HEC resources.

For accounting, pricing, and billing individual solutions mostly based on flatrates have been used (DC and HPC). Future focus will concentrate on compound units and modular solutions. An integrated distributed accounting and billing solution considering national and international legal regulations must be implemented for infonomics purposes.

Up to now various methods and technologies have been implemented for use with the new features. There will have to be strong standardisation efforts for secure and ergonomic access, quality management, and new types of content specification, allowing flexible separation of data, information, and functional parts. Considering legal aspects is crucial for the design of these specifications in order to create a suitable structure based on the legal frameworks.

Physical aspects for infonomics systems have often been neglected in the past. Data sizes have been sized small and only transferred for small distances. Prominent topics are broadband for public use with data transfer and physical identification for real world tracking. These will be needed for transfer of large amounts of computation data in national and international context, for monitoring and logistics in exploration and environmental management. Projects are experimenting with large data sizes ($>$ 100 TeraBytes) on international data transfer for use with applications. These are expected to be industry topics in the near future. Broadband networks will allow to transfer larger amounts of data using secure channels on external networks.

The implemented data handling is suitable for data types and applications currently used. The case studies of the last years have shown that there is need for standardised cognostic categories. Large size applications with composite data types do need data-categorisation for generalising, integration, automation, georeferencing, and search engine facilities in order to minimising conflicts with legal regulations.

Operation and updates will be interesting for providers and industry. Legal regulations demand a transparent least-invasive access and update concept for complex information and computing systems.

A. Primary topics on technology and legal issues

A number of primary topics resulting from the interactions within the columns of the infonomics system exist from user point of view: security, safety/privacy, consistency, international standards, legal issues, and primary associated laws and regulations, identified with the GEXI case studies (Table III). This table shows some of the most important components, that are not worked out ultimatively

Table III
RESULTING PRIMARY TOPICS AND FUTURE EMPHASIS.

Topic	*Column*	*Sec.*	*Saf.*	*Con.*	*Int.*	*Leg.*	*Law (DE)*
			Services				
Services	DG	(✓)	(✓)	—	—	—	BDSG GeoZG
Georeferencing	DG	—	—	—	—	✗	BDSG
Automation	DG	—	—	—	(—)	✗	BDSG
			Communication / Transfer				
Networking	DH	(✓)	(✓)	(✓)	(—)	(—)	BDSG
			Distributed Computing				
Accounting	DHG	(✓)	(✓)	(✓)	✗	✗	BDSG
Billing	DHG	(✓)	(✓)	(✓)	✗	✗	BDSG
			High Performance Computing				
Computing	HD	✗	✗	✗	✗	✗	BDSG UrhG
Networking/IB	HD	✗	✗	✗	✗	✗	BDSG UrhG
Storage	DH	(✓)	✗	✓	✗	✗	BDSG UrhG
			Disciplines				
Inf. Systems	GD	(✓)	(—)	(✓)	✗	✗	IFG GeoZG IWG
Geosciences	G	(—)	(—)	(—)	✗	✗	IFG GeoZG IWG
Exploration	OG	(—)	(—)	(—)	✗	✗	IFG UrhG IWG
Environment	OG	(—)	(—)	(—)	✗	✗	IFG UIG IWG
Medicine	OG	(—)	(—)	(—)	✗	✗	IFG BDSG IWG
e-Science	OGD	(—)	(—)	(—)	✗	✗	IFG BDSG IWG

(✓ partially done, ✗ worked on within interdisciplinary cooperations, G: Geo, H: HPC, D: DC, O: other). Only some topics like storage consistency can be considered done for midterm, as there are means of creating suitable solutions. Other topics like safety within the productive employment of geosciences, exploration, medicine, and e-Science can be considered specific to their discipline. There are some topics most important as they concern several of the topics and disciplines: facilitate market use of HEC resources and creating collaboration frameworks for national to international use.

B. Collaboration portal

The GEXI case studies have shown the additional need for a single access point, a frontend portal. Resulting from the integration of the consolidated academic and industrial interests, Figure 4 shows a sketch of the prototype GEXI portal addressing the geo-exploration-energy market.

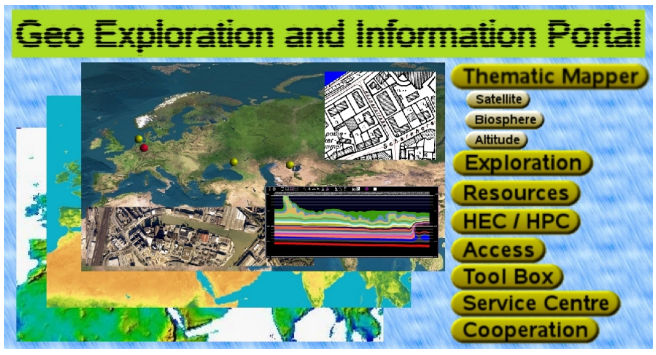


Figure 4. Sketch of the GEXI portal.

The portal is aimed to provide a collective collaboration

platform for the columns of the infonomics system in order to prove the conjoint economic and legal feasibility. Clients for disseminating the provisioned services, information and distributed resources will be integrated with this portal. Such a portal is capable of providing a solution to technical-legal diversities as with access, services accounting, licensing, QoD, automation, personalisation, security, and internationalisation, handling focus points within and between the collaborating columns at the backend and reducing complexity.

IX. Implementation case studies

A. Application scenarios

The information and computing system components make use of various technologies, IPC, sandboxing, embedded applications, browser plugins, remote execution, network protocols, computing interfaces as well as public and sensible data. Figure 5 shows some of the basic application scenarios.

There exists a number of scenarios showing how "trust in computation" and "trust in information" can more easily be achieved by reducing complexity for the partners in otherwise very complex systems. The following sections give an extended presentation of selected case studies, implemented following the collaboration framework. The focus is on implementation, legal issues, and security aspects faced by the resources, services, and disciplines columns. The following collaboration matrices show what topics the columns Resources (R), Services (S), and Disciplines (D) had to take care while realising the components regarding

- implementation (i),
- legal issues (l),
- security (s).

The column partners have been responsible for the topics respecting the work packages designed respecting the collaboration framework.

B. Environmental information and computing

Various information resources are available for environmental and energy exploration. Mostly all of the implementations making use of these resources are standalone systems. Computing resources are not considered part of the implementations at all.

For the future, integrating information, monitoring, management, and computing systems is necessary for effectively and efficiently using these resources. As an example, the information on private, governmental, and industrial land use, national parks information, and energy and mineral resources exploration is very complex. The parameters of wind energy and solar energy are highly dynamical. Calculation of weather impacts, construction and simulation of new facilities offshore and onshore leads to new demands on information and computing.

The case study showed that "trust in computation", reliability and suitability of information, QoD, and security of critical investments are most important for the academic and industrial partners. As a result, for this scenario it has been regarded necessary for the services to implement and configure a complex combination of the following features:

- dedicated networks,
- firewalls, access lists, routing,
- sandboxes,
- trusted scripting,
- shared and non-shared use,
- queue limits,
- demilitarised zone,
- access control and keys,
- resources monitoring and accounting,
- local auditing, communication packet filtering, security management,
- job logging,
- encrypted data transfer and communication,
- services monitoring,
- on-site support and management.

Table IV gives a summarising excerpt of the collaboration matrix, showing columns (Resources, Services, Disciplines), topics (implementation, legal issues, security), the location where the activity is concentrating on, and the focus topics with their priority (highest priority is marked '!!!').

As the table shows, emphasised priority with this case is on computational aspects. No individual client applications have been regarded necessary for the users with this case study. The standard client is a Secure Shell client accessing resources and implementing individual automation facilities using access key pairs and batch system services. User groups will develop many of the algorithms and tools needed for processing. Public keys are not the primary means for authentication with this scenario. Data encryption is triggered by users. Decision flow and overall priority is on the users side. Security and reliability concentrates on computation and shared resources usage. Legal aspects are handled on disciplines/users side. The collaboration matrix shows that both the disciplines and services have had to concentrate on the legal topic-framework implementation as the services layer had to be staffed with privileged users as well as operated and managed by services groups.

With assignment of distributed resources, legal and security aspects of information security and access require transparent handling on services side. The Active Source concept has been used for the implementation, based on a modular server-client and services interfaces architecture. Figure 6 shows a screenshot from the environmental application case study, showing an Active Map of the National Parks information system and some distributed online information [24], [25]. The implementation targets on environment-geosciences information and energy exploration based on High Performance Computing (HPC) and Distributed Computing (DC) services.

For the current collaboration, [26] resources management

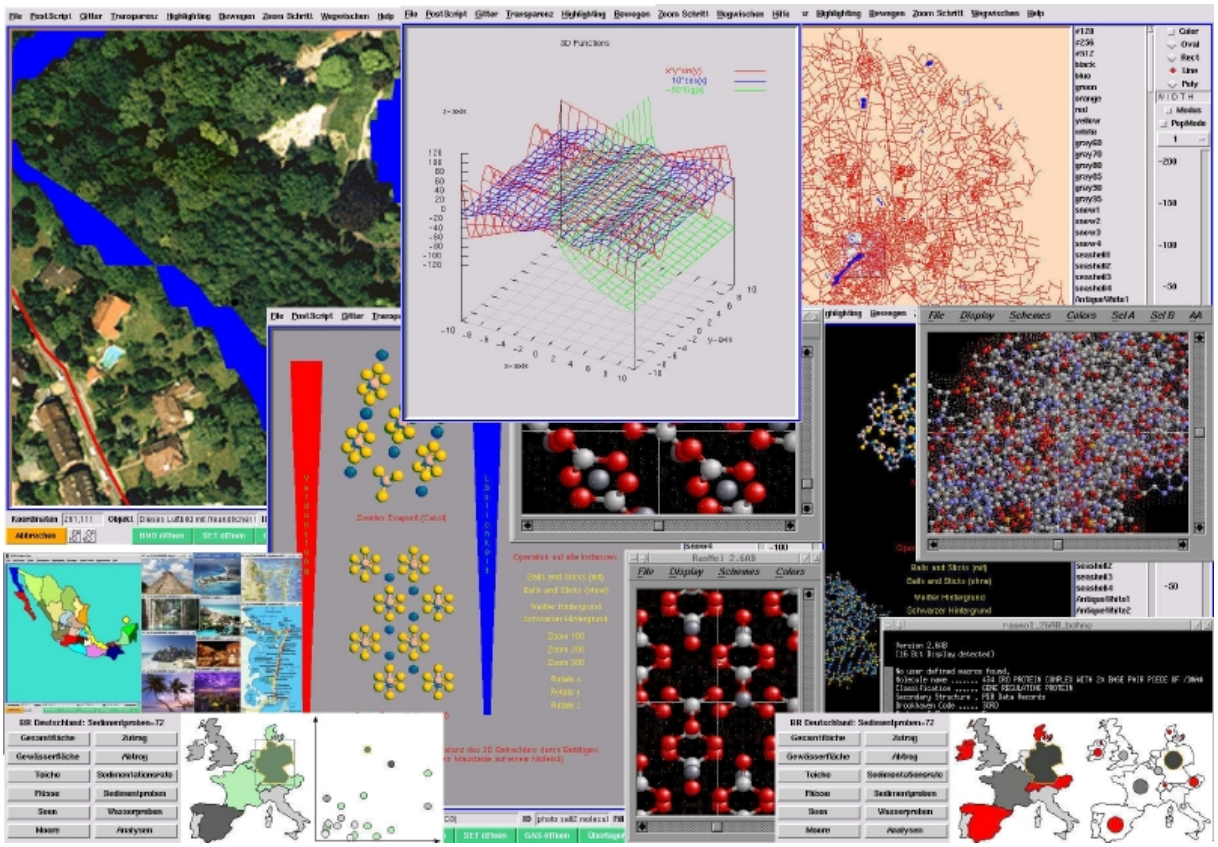


Figure 5. Basic application scenario implementations from the GEXI case studies examples.

for geosciences, energy sciences, and mobility is based on the collaboration framework. A couple of aspects of future Enterprise Resource Planning (ERP) systems are most relevant for supporting integrated HEC and Information Systems in industrial applications in this scenario:

- combining distributed and parallel techniques,
- transaction processing,
- highly parallel communications,
- loosely parallel communications of highly parallel applications,
- industrial systems,
- expanding scalability,
- improving the number of compute nodes for industry fields of application, currently less than 100 compute nodes,
- open up access to scientific scenarios with currently up to over 500 compute nodes,
- preparing new solutions for complex industrial information-computing systems.

These are objects for further implementation on the resources and services columns. The implementations for global monitoring software and expanding the physical distribution for PSI strongly depend on their availability.

The case study further showed that “trust in computing” and reliability are most important for the academic and industrial partners. It has been possible to transparently separate nearly all of the implementation aspects for the three columns. The most prominent conjoint implementation issues having to be worked on for the future is national laws and regulations as far as there is no general solution for securing critical data, and computation with distributed usage cannot be supported by signing some kind of black-box “computation objects”.

C. Epidemology information and computing

Interdisciplinary research in the ecology and epidemology of vector-borne diseases produces huge amounts of data regarding to biological and epidemological processes [27], [28].

In epidemology a vector is an insect or any living carrier that transmits an infectious agent. Examples are hematophagous arthropod vectors such as mosquitos, ticks and flies which are responsible for transmitting protozoa, bacteria, and viruses between vertebrate hosts, causing diseases as Malaria, Lyme disease, and Sandfly Fever. Examples for processes are disease prevalence, abundance, and distribution of living organisms.

These parameters are highly dynamical and are influenced in time and space by a number of biotic (e.g. vegetation) and abiotic (e.g. temperature and humidity) factors.

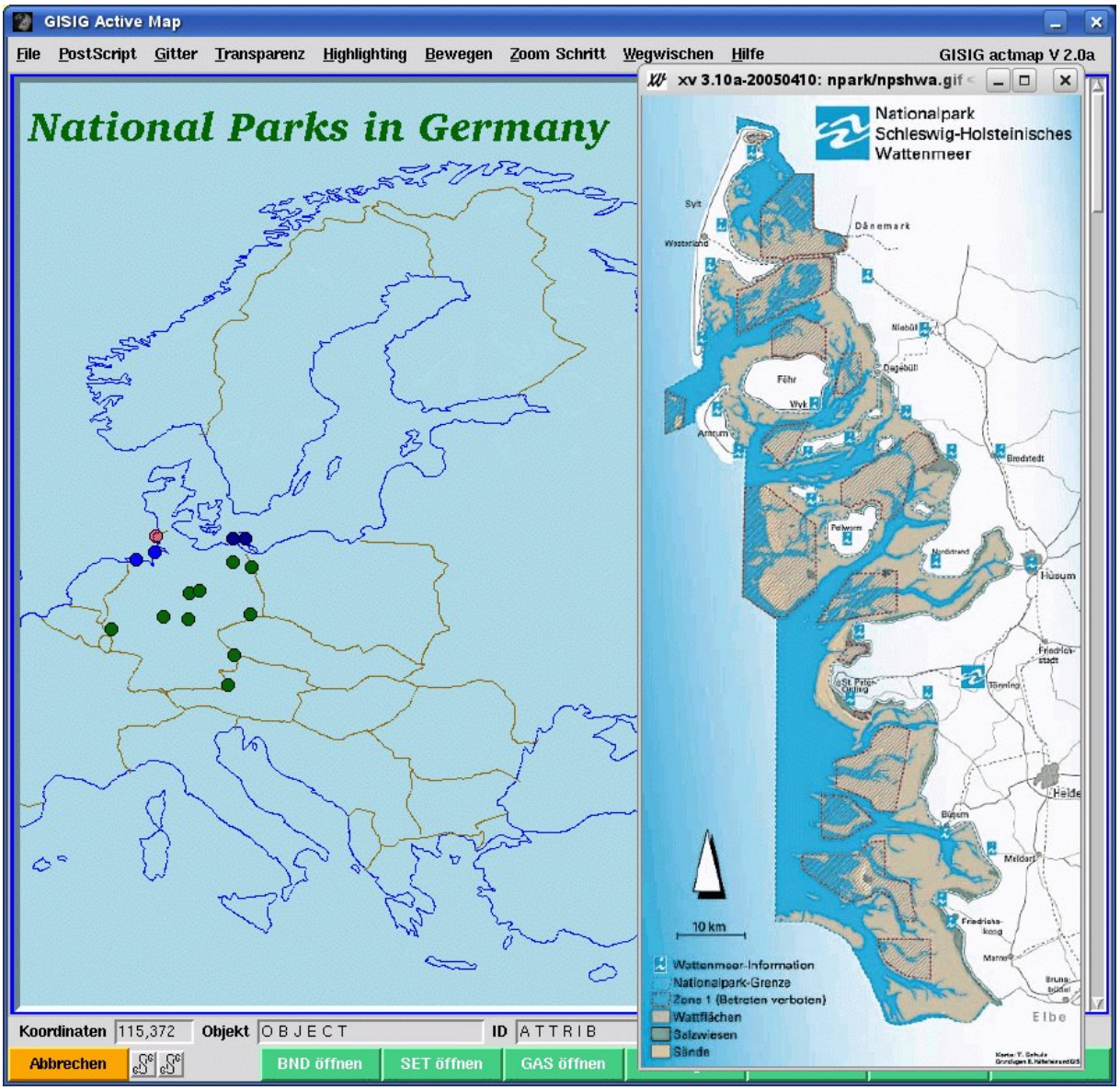


Figure 6. Environmental application case study showing a distributed multi-discipline Active Source implementation.

There are three main types of objects: epidemological objects, environmental objects, and sociocultural objects. The following listing showing the simplified example matrix.

- epidemological: disease case numbers, vector distribution, number of vectors with pathogen,
- environmental: landscape structure, meteorological conditions, vegetation coverage, micro climate change,
- sociocultural: interchange of organisms by human travel and commerce.

The case study showed that using security enhanced, PKI/CA and PKC [29] based integrated Information and Computing System components for scientific application including industry participation has been regarded most suitable. By these means evaluation of data, generating various views, querying distributed information and visually summarising information by calculating multi-dimensional views varying in time and spatial representation can be supported in a secure and flexible way. For locally distributed information spatial information processing algorithms can be employed. Partners from several disciplines will be able to evaluate and analyse this data, e.g., human health epidemologists, public health authorities, and physicians.

A well balanced infonomics system will allow for the input, storage, data manipulation, analysis, and visual presentation of georeferenced data and is particulary suitable for identifying local clusters of diseases, and for analysing spatial relationships between diseases and risk factors.

Table V gives a summary for the collaboration matrix. As the table shows, emphasised priority with this case is on content and utilisation aspects. Special client applications

Table IV
COLLABORATION MATRIX FOR THE ENVIRONMENTAL INFORMATION AND COMPUTING CASE STUDY.

Column	*Topic*	*Location*	*Focus : Priority*
		Physical	
R – –	i – s	provider	compute resources security : !!
R – –	i – s	provider	storage resources security : !!
		Application	
– – D	i – s	prov./user	content access client : !
– – D	i – s	user	services automation : !
– S –	i – s	provider	process communication security : !!!
– S –	i – s	user	execution security : !!!
		Computation	
R – –	i – s	prov./user	comp. res. availability : !!!
R – –	i – s	prov./user	resources access security : !!!
R – –	i – s	provider	interfaces security : !!!
R – –	i – s	provider	computation reliability : !!!
– S –	i – s	provider	power-on encryption : !!!
		Content Information	
– – D	– l –	user	pollution data legal regulations : !!
– – D	i – s	user	functional content security : !!
– – D	i – –	user	object data regulations : !!
– – D	i – s	user	event data security : !!
– – D	i – s	user	autoevent data security : !!
– – D	i – s	user	power-off encryption : !!
– – D	- l –	user	national legal context : !!
		Utilisation Information	
– S –	i – s	provider	content data transfer : !
– S –	i – –	provider	user and client access : !
– S –	i – s	provider	content storage security : !
– S –	i – –	provider	modification and transfer : !
– S –	i – –	provider	reliability of communication : !
– – D	l – –	user	information signing : !
– – D	i – –	user	distribution of information : !
– S D	– l –	provider	national laws and regulations : !

Table V
COLLABORATION MATRIX FOR THE EPIDEMOLOGY INFORMATION AND COMPUTING CASE STUDY.

Column	*Topic*	*Location*	*Focus : Priority*
		Physical	
R – –	i – s	provider	compute resources security : !!
R – –	i – s	provider	storage resources security : !!
		Application	
– S –	i – s	prov./user	content access client : !!!
– S –	i – s	prov./user	services automation : !!!
– S –	i – s	provider	process communication security : !!!
– S –	i – s	user	execution security : !!!
		Computation	
R – –	i – s	prov./user	comp. res. availability : !
R – –	i – s	prov./user	resources access security : !!
R – –	i – s	provider	interfaces security : !!
R – –	i – s	provider	computation reliability : !
– S –	i – s	provider	power-on encryption : !
		Content Information	
– – D	– l –	user	epidemiol. data legal regulations : !!!
– – D	i – s	user	functional content security : !!
– – D	i – –	user	object data regulations : !!
– – D	– l s	user	content data security : !!!
– – D	i – s	user	autoevent data security : !!
– – D	i – s	user	power-off encryption : !!
– – D	- l –	user	national legal context : !!
		Utilisation Information	
– S –	i – s	provider	content data transfer : !!!
– S –	i – –	provider	user and client access : !!!
– S –	i – s	provider	content storage security : !!!
– S –	i – –	provider	modification and transfer : !
– S –	i – –	provider	reliability of communication : !
– – D	l – –	user	information signing : !!!
– – D	i – –	user	distribution of information : !
– – D	– l –	provider	national laws and regulations : !

are necessary for authors and users with this case study. The user groups will not develop services on their own, neither algorithms nor tools needed for processing. The centre of the information system is the PKI/CA and PMI/AA infrastructure. Processing uses the signed objects within the services layer, only accessible via dedicated service interfaces.

Security therefore concentrates on information and more or less computation. The physical shared resources usage can be critical due to the storage/scheduling location being currently not predictable while being effective. The services column is responsible for the system and client security.

Legal aspects are handled on disciplines side. The collaboration matrix shows that the disciplines column has to concentrate on the legal topic-framework implementation.

With this scenario "trust in information" is twofold, regarding the content information domain and the utilisation information domain. It has been possible to transparently separate nearly all of the implementation aspects for the three columns.

The case study showed that for the application within the integrated information and computing system the three main types of objects need role-based data access for users and clients. Objects have to be signed with digital signature and timestamps of the originating authors and manipulation. For real life scenarios network transfer encryption has to be used. Due to a distributed storage environment host-side encryption has been regarded necessary.

Figure 7 shows the workpackage layers view. As well as in other disciplines, in epidemiology sciences on the one hand there is a strong need to assure accurate data objects all over the life-cycle of objects, thus for content guaranteeing "trust in information". On the other hand a suitable access control infrastructure has to be established.

Strong authentication and authorisation by means of cryptographic techniques specified as Public Key Certificates (PKC) and Attribute Certificates (AC) in Public Key Infrastructure (PKI) and Privilege Management Infrastructure (PMI) environments [30] provides a framework for adressing important security considerations of authentication, confidentiality, authorisation, and integrity (PKI) and allows for particularly controlled data access (PMI). In this regard the authority (CA) signs the public user keys in order to maintain the integrity of the public key, expiration information

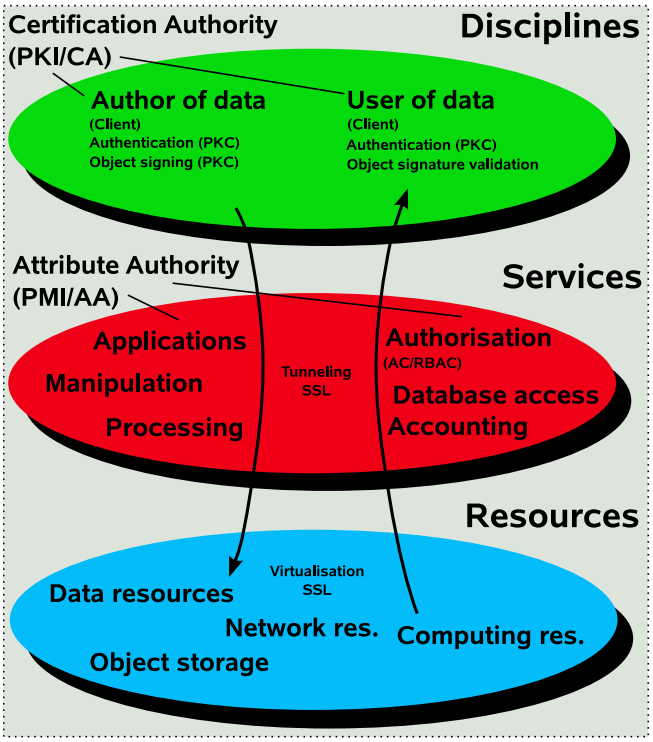


Figure 7. Workpackage layers for the epidemology information and computing case study development.

and other important information contained within the user certificates. Attribute Authority (AA) is the authority which assigns attributes (permissions and privileges) by signing the attribute certificate. Digital signatures are used in both PKI and PMI as the mechanism which binds the issueing certification authority to the certificate. CA and AA are separate authorities and should be established independently.

A sophisticated trust management system using X.509 Attribute Certificates [31] can be used to store the user roles, based on Role Based Access Control (RBAC) [32].

In case of multidisciplinary working and development groups (Figure 7), intending epidemiological analysis of infectious diseases, the procedure has been used as follows. For an operation example with the case study this means:

- **Disciplines layer:** The author of a data object, the originator, e.g. co-worker of a human health organisation, signs the created object, e.g. disease case numbers, with the private key of the author and according timestamps.
- **Services layer:** The object is processed by the application services.
- **Resources layer:** The processed object is stored to the distributed storage.
- **Disciplines layer:** An user, e.g. a member of a research team at an university, requests an object.
- **Services layer:** The user is authenticating at the authentication service via AC/RBAC, the signature of the corresponding author is validated. The authentication service requests the object or service operation.
- **Resources layer:** The object is collected from the distributed resources.
- **Services layer:** The object is calculated, accounted, and provisioned via services for the user client.

X. Lessons Learned

The environmental case has been found more heterogeneous than the epidemological example. This is the result of the fact that in these disciplines many interest groups are dealing with algorithms, simulations and spatial planning for over the last decades. These disciplines will be able to develop some own services components and make broader use of sophisticated High End Computing resources. It has been more transparent to define interfaces for discussing and integrating legal frameworks and regulations into the multi-disciplinary implementation process. Nevertheless for the majority of use cases, smaller scenarios can be seen where strict separation of disciplinary work, services and development, and operation of resources will be lived.

This leads to the conclusion that in the future of integrated information and computing systems we will need to create means of securely submitting modular application components into the services pipeline.

A collaboration framework for development and operation in combination with and integrated information and computing systems handling these features should be able to address many of the heterogeneous conditions regarding implementation, legal issues, and security existing in national as well as international context. It will support stearing of data, information, and application workflows being conform with legal regulations and data security standards as well as obeying policies.

XI. Conclusion and Future Work

Developing international cooperations within the fields of geosciences, exploration, and computing, based on an interdisciplinary collaboration framework is regarded a perfect solution for all partners in the GEXI study for the different case study scenarios, in order to modularise information system development and reduce complexity. Technology suitable for solving open problems with implementation, legal environment or trust is still in the genesis. For instance with trust and encryption regarding Distributed Computing and shared resources, full homomorphic encryption techniques are desirably. Currently basic algorithms are available.

The modular integration of services and disciplines within a collaboration framework has proven best results to be flexible and efficient for large international projects with various legal characteristics as separating technical and legal work packages. With this, the future allocation of responsibilities and integration of specialist frameworks has become a more transparent process.

The ongoing work is oriented towards the consequences resulting from the evaluation. The legal aspects for these topics will attend the next steps in order to elaborate the framework for use with a collaboration portal with support from legal working groups.

At this point it is essential for complex implementations to integrate the national legal regulations (in Germany e.g. IFG, GeoZG, IWG) with securely managing content and workflows into the international context as well as to refine the interactions within the collaborating sections:

1) For legal issues with geosciences and exploration, the aspects of data contents, cognostics, data combination and automation (georeferencing) and parallelisation for HPC and shared resource usage are in the focus. Integration with the national and international legal frameworks (e.g. GDI-DE, INSPIRE, GSDI) will have to be forced in order to accomplish a base for future information and decision making systems for commercial and educational purposes.
2) Regarding the Distributed Computing and services column, desktop user interfaces, services and security, networking, and undisputable distributed accounting have been set top on the working list.
3) With High Performance Computing and resources standardisation and a secure networking model with PKI, privacy, and encryption support for future HEC architectures development is priority.

Further on with implementation and legal issues, the security aspect are on the rise for any complex system. Even though PKI technology offers means to attest, identify, manage the exchange of encryption keys an secure transmission between parties, there has not been broad-based adoption of PKI technology by public and private organisation. After all, a significant number of countries recognise digital signatures as legally binding. In case of security enhanced integrated information and computing system components object signing provides a robust solution to facilitate "trust in information" and to overall support "trust in computing". In order to put this implementation into international public practice there is a need for future PKI development and deployment offering a global public key cryptosystem for the Future Internet.

Preliminary work has created a common base for an ethical understanding of cross-disciplinary use of data. Various trust situations, important for services providers on the one hand as well as for disciplines on the other hand could be described, handled, and implemented with the separation of work packages. This work showed that it is possible to bring complex information and computing systems to life, being able to create interfaces that can also be interfaces between the logical columns and interest groups.

ACKNOWLEDGEMENTS

We are grateful to all national and international academic and industry partners in the GEXI cooperations for the innovative constructive work and especially for the environmental and energy exploration case study input, security advisor and representative Mrs. Birgit Gersbeck-Schierholz for implementing and coordinating the epidemology disciplines case study, and to the colleagues at the Leibniz Universität Hannover, at the IRI, the North-German Supercomputing Alliance (HLRN), WWU, ZIV, D-Grid and the participants of the postgraduate European Legal Informatics Study Programme (EULISP) for prolific discussion of scientific, legal, and technical aspects as well as to the staff at ZIV, ZIB, L3S and associated HPC companies for supporting this work by managing and providing HEC resources over the years.

REFERENCES

[1] C.-P. Rückemann, "Legal Issues Regarding Distributed and High Performance Computing in Geosciences and Exploration," in *Proceedings of the International Conference on Digital Society (ICDS 2010), The International Conference on Technical and Legal Aspects of the e-Society (CYBERLAWS 2010), February 10–16, 2010, St. Maarten, Netherlands Antilles / DigitalWorld 2010, International Academy, Research, and Industry Association (IARIA)*. IEEE Computer Society Press, IEEE Xplore Digital Library, 2010, pp. 339–344, Berntzen, L., Bodendorf, F., Lawrence, E., Perry, M., Smedberg, Å. (eds.), ISBN: 978-0-7695-3953-9, URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5432414 (PDF) [accessed: 2010-12-26], (Best Paper Award).

[2] "Geo Exploration and Information (GEXI)," 1996, 1999, 2010, URL: http://www.user.uni-hannover.de/cpr/x/rprojs/en/index.html#GEXI (Information) [accessed: 2011-01-01].

[3] C.-P. Rückemann, *Accounting and Billing in Computing Environments*. Business Science Reference, IGI Global, Hershey, Pennsylvania, USA, Oct. 2009, 25 pages, Chapter X, in Pankowska, M. (ed.), Infonomics for Distributed Business and Decision-Making Environments: Creating Information System Ecology, 421 pages, ISBN: 978-1-60566-890-1, URL: http://www.igi-global.com/reference/details.asp?ID=34799 (Information) [accessed: 2010-12-26].

[4] C.-P. Rückemann, "Using Parallel MultiCore and HPC Systems for Dynamical Visualisation," in *Proceedings of the International Conference on Advanced Geographic Information Systems & Web Services (GEOWS 2009), February 1–7, 2009, Cancun, Mexico / DigitalWorld 2009, International Academy, Research, and Industry Association (IARIA)*. IEEE Computer Society Press, IEEE Xplore Digital Library, 2009, pp. 13–18, Dragicevic, S., Roman, D., Tanasescu, V. (eds.), ISBN: 978-0-7695-3527-2, URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4782685&isnumber=4782675 (PDF) [accessed: 2010-12-26], (Best Paper Award).

[5] C.-P. Rückemann, "Beitrag zur Realisierung portabler Komponenten für Geoinformationssysteme. Ein Konzept zur ereignisgesteuerten und dynamischen Visualisierung und Aufbereitung geowissenschaftlicher Daten," Dissertation, Westfälische Wilhelms-Universität, Münster, Deutschland, 2001, 161 (xxii + 139) Seiten, Ill., graph. Darst.,

Kt., URL: http://wwwmath.uni-muenster.de/cs/u/ruckema/x/dis/download/dis3acro.pdf [accessed: 2010-12-26].

[6] C.-P. Rückemann, "Geographic Grid-Computing and HPC empowering Dynamical Visualisation for Geoscientific Information Systems," in *Proceedings of the 4th International Conference on Grid Service Engineering and Management (GSEM), September 25–26, 2007, Leipzig, Deutschland, co-located with Software, Agents and services for Business, Research, and E-sciences (SABRE 2007)*, R. Kowalczyk, Ed., vol. 117. GI-Edition, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik e.V. (GI), 2007, pp. 66–80, ISBN: 78-3-8579-211-6, ISSN: 1617-5468.

[7] C.-P. Rückemann, "Dynamical Parallel Applications on Distributed and HPC Systems," *International Journal on Advances in Software*, vol. 2, no. 2, 2009, ISSN: 1942-2628, URL: http://www.iariajournals.org/software/ [accessed: 2010-12-26].

[8] "Global Monitoring for the Environment and Security (GMES)," 2010, URL: http://www.gmes.info [accessed: 2011-01-01].

[9] "Global Earth Observation System of Systems (GEOSS)," 2010, URL: http://www.earthobservations.org/geoss.shtml [accessed: 2011-01-01].

[10] "Group on Earth Observations (GEO)," 2009, URL: http://www.earthobservations.org [accessed: 2011-01-01].

[11] "Shared Environmental Information System (SEIS)," 2010, URL: http://ec.europa.eu/environment/seis/ [accessed: 2011-01-01].

[12] "Global Spatial Data Infrastructure (GSDI)," 2010, URL: http://www.gsdi.org [accessed: 2011-01-01].

[13] "INfrastructure for SPatial Information in Europe (INSPIRE)," 2010, URL: http://www.ec-gis.org/inspire [accessed: 2011-01-01].

[14] "Geodateninfrastruktur Deutschland (GDI-DE)," 2010, URL: http://www.gdi-de.org [accessed: 2010-12-26].

[15] "European Public Sector Information (EPSI)," 2010, URL: http://www.epsiplus.net [accessed: 2011-01-01].

[16] C.-P. Rückemann, *Fundamental Aspects of Information Science, Security, and Computing (Lecture)*. EULISP Lecture Notes, European Legal Informatics Study Programme, Institute for Legal Informatics, Leibniz Universität Hannover (IRI / LUH), 2009, URL: http://www.eulisp.de [accessed: 2011-01-01].

[17] "SGAS – SweGrid Accounting System," 2010, URL: http://www.sgas.se/ [accessed: 2011-01-01].

[18] P. Gardfjäll, E. Elmroth, L. Johnsson, O. Mulmo, and T. Sandholm, "Scalable Grid-wide Capacity Allocation with the SweGrid Accounting System (SGAS)," *Concurrency and Computation: Practice and Experience*, vol. 20, no. 18, pp. 2089–2122, 2008, URL: http://www.cs.umu.se/~elmroth/papers/sgas_revised_aug_2007.pdf [accessed: 2010-12-26] (Preprint).

[19] T. Sandholm, P. Gardfjäll, E. Elmroth, L. Johnsson, and O. Mulmo, "A Service-oriented Approach to Enforce Grid Resource Allocations," *International Journal of Cooperative Information Systems*, vol. 15, no. 3, pp. 439–459, 2006, URL: http://www.cs.umu.se/~elmroth/papers/SGASIJCIS_2006.pdf [accessed: 2010-12-26].

[20] E. Elmroth and P. Gardfjäll, "Design and Evaluation of a Decentralized System for Grid-wide Fairshare Scheduling," *e-Science 2005, First IEEE Conference on e-Science and Grid Computing*, pp. 221–229, 2005, URL: http://www.cs.umu.se/~elmroth/papers/fsgrid.pdf [accessed: 2010-12-26].

[21] E. Elmroth, P. Gardfjäll, O. Mulmo, and T. Sandholm, "An OGSA-based Bank Service for Grid Accounting Systems," *State-of-the-art in Scientific Computing*, vol. 3732, pp. 1051–1060, 2006, URL: http://www.cs.umu.se/~elmroth/papers/egms_para04.pdf [accessed: 2010-12-26].

[22] E. Elmroth and D. Henriksson, "Distributed Usage Logging for Federated Grids / Future Generation Computer Systems," *The International Journal of Grid Computing: Theory, Methods and Applications*, submitted 2009.

[23] E. Elmroth, F. Galán, D. Henriksson, and D. Perales, "Accounting and Billing for Federated Cloud Infrastructures," *In: Proceedings of the Eighth International Conference on Grid and Cooperative Computing (GCC 2009), J. E. Guerrero (ed.)*, pp. 268–275, 2009, URL: http://www.cs.umu.se/~elmroth/papers/eghp_gcc2009.pdf [accessed: 2010-12-26].

[24] "Nationalpark Schleswig-Holsteinisches Wattenmeer," 2010, URL: http://www.sh-nordsee.de/nationalpark/ [accessed: 2010-12-26].

[25] "National Parks, Germany," 2010, URL: http://www.nationalparke.de/ [accessed: 2010-12-26].

[26] C.-P. Rückemann, "Future Geo-Exploration Information and Computing Systems Created by Academia-Industry Collaboration," in *Proceedings of the 9th International Scientific-Practical Conference 2010 (HTFR 2010), April 22–23, 2010, Saint Petersburg, Russia*. Saint Petersburg: Saint Petersburg University Press, 2010, pp. 254–258, ISBN: 978-5-7422-2558-4, URL: http://htfr.org [accessed: 2010-12-26].

[27] C. G. Moore, "Interdisciplinary research in the ecology of vector-borne diseases: Opportunities and needs," *Journal of Vector Ecology*, vol. 33, no. 2, pp. 218–224, 2008, DOI: 10.3376/1081-1710-33.2.218, URL: http://www.sove.org/Journal%20PDF/December%202008/1-Moore%2008-56.pdf [accessed: 2010-12-26].

[28] S. I. Hay, A. Graham, and D. J. Rogers, Eds., *Global Mapping of Infectious Diseases: Methods, Examples and Emerging Applications*. Academic Press, Edition: Pap/Dvdr R, 2007, iSBN-10: 0-12-031764-8, ISBN-13: 978-0-12-031764-6, ISBN-13: 978-0-12-031766-0 (DVD), ISBN-10: 0-12-031766-4 (DVD).

[29] B. F. S. Gersbeck-Schierholz, "Trustworthy Communication by Means of Public Key Cryptography," 2010, URL: http://www.rrzn.uni-hannover.de/fileadmin/it_sicherheit/pdf/pki2010_gersbeck.pdf [accessed: 2010-12-26].

[30] "ITU-T Recommendation X.509 ISO/IEC 9594-8, The Directory: Authentication Framework," 2000, URL: http://wwwitu.int/itu-t/recommendations [accessed: 2010-12-26].

[31] D. W. Chadwick and A. Otenko, "The PERMIS X.509 role based privilege management infrastructure," *Future Generation Computer Systems (FGCS)*, vol. 19, pp. 277–289, 2003, DOI: 10.1016/S0167-739X(02)00153-X, URL: http://portal.acm.org/citation.cfm?id=770786 [accessed: 2010-12-26].

[32] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-Based Access Control Models," *IEEE Computer*, vol. 29, no. 2, pp. 38–47, 1996, DOI: 10.1109/2.485845, ISSN: 0018-9162, URL: http://csrc.nist.gov/rbac/sandhu96.pdf [accessed: 2010-12-26].

ASPF: A Policy Administration Framework for Self-Protection of Large-Scale Systems

Ruan He Marc Lacoste
Orange Labs
Security and Trusted Transactions Dept.
{ruan.he,marc.lacoste}@orange-ftgroup.com

Jean Leneutre
Telecom ParisTech
Network, Mobility and Security Dept.
jean.leneutre@telecom-paristech.com

***Abstract*—Despite its potential to tackle many security challenges of large-scale systems such as pervasive networks, self-managed protection has been little explored. This paper addresses the problem from a policy management perspective by presenting a policy-driven framework for self-protection of pervasive systems called ASPF (Autonomic Security Policy Framework). Enforced authorization policies in a device are adapted according to the security context, both at the network and device levels. ASPF describes how an autonomic security manager may control OS-level authorization mechanisms supporting multiple classes of policies. Evaluation of an ASPF implementation shows that the framework enables effective self-protection of pervasive systems. ASPF is also applicable for autonomic security management of other types of large-scale infrastructures such as cloud environments.**

***Keywords*-Autonomic Computing, Self-Protection, Policy Management, Authorization, Pervasive Networks.**

I. INTRODUCTION

Advances in pervasive networking are rapidly taking us to the final frontier in security, revealing a whole new landscape of threats. In open and dynamic environments, malicious nodes may enter a network undetected, and various malwares may invisibly install themselves on a device. When roaming between heterogeneous networks, each with its own protection requirements, a device may also take advantage of security policy conflicts to gain unauthorized privileges. In embedded settings including limited and often unstable computing and networking resources, denial of service attacks are easier, with little lightweight security countermeasures. Finally, these decentralized, large-scale systems make end-to-end security supervision difficult. Administration by hand is clearly impossible, with the risk of some sub-system security policies not being up-to-date. These threats may only be mitigated with mechanisms highly adaptable to execution conditions and security requirements (e.g., supporting multiple authorization policies), with limited overhead. Above all, protection mechanisms should be self-managed [1], following the autonomic approach to security introduced by IBM [2], which defines a *self-protecting system* as a system that "can anticipate, detect, identify and protect [itself] against threats." [3].

To realize context-aware autonomic adaptations, the policy-driven paradigm has successfully demonstrated its flexibility and generality [4]: system functionalities are governed by a set of policies. As the context changes, other policies may be selected to activate within the system functions better adapted to its new environment. Unfortunately, this type of design was little applied to self-protection of pervasive systems.

In this paper, we validate the viability of this approach by presenting a policy-driven security management framework called *ASPF (Autonomic Security Policy Framework)*. ASPF describes the design of an autonomic security manager for pervasive systems. The framework is built on an earlier implemented OS security architecture called *Virtual Security Kernel (VSK)* [5]–[7] that specifies the managed security mechanisms. VSK implements kernel-level policy-neutral authorization, and supports dynamic policy reconfiguration, but without describing any control strategy of adaptation.

The original features of this framework are the following:

- ASPF enables the selection of the most appropriate authorization policy to be enforced in the device in order to match the estimated risk level of the current environment. Two levels of adaptation are possible, policies being tuned (or generated) according to the security context of the network and of the device.
- Policies are specified in an XACML extension for the attribute-based model of access control [8], which provides a fairly generic manner to describe permissions in open systems.
- An authorization architecture is also defined to refine the ASPF models, and is implemented above the VSK authorization mechanisms.

Performance, resilience, and security evaluation results show that the combined ASPF and VSK frameworks enable to achieve effective self-protection (Section IX-B evaluates the autonomic maturity level achieved with ASPF regarding security mechanisms). Moreover, ASPF is generic enough to be applied to other types of large-scale infrastructures such as cloud computing environments by defining the proper framework refinement.

This paper is organized as follows. After reviewing related work (Section II), we introduce briefly our self-protection architecture (Section III). We then describe the ASPF design principles (Section IV), policy model (Section V), framework (Section VI) and authorization architecture (Section VII). We present an ASPF implementation over the VSK mechanisms (Section VIII), and some evaluation results (Section IX). We finally show how ASPF may be refined for self-protection in cloud environments (Section X).

II. Related Work

Self-protection has so far been explored very little. While quite an early idea [2], it was discussed at the level of principles with few frameworks available, mainly for enterprise information systems [9], [10]. To orchestrate the components needed for autonomic security management, a policy-driven design [4], [10] seems promising, since the approach has been successfully applied to other self-* properties: indeed, several generic policy management frameworks [11]–[13] have been proposed to automate device and network reconfigurations to respond to context changes. Unfortunately, these frameworks hardly considered security. Notable exceptions are [13] for large organizations and [14] for pervasive systems which supports authorization and obligation policies. But with those frameworks, it remains unclear how to specify and federate authorization policies described in different security models to overcome heterogeneity of network security policies.

Three main elements seem to be missing: (1) descriptions of self-protection strategies; (2) specifications of security policies; and (3) authorization mechanisms supporting multiple policies and/or their reconfiguration. A promising approach for (1) is based on Domain-Specific Languages (DSLs) [15], but does not yet address security. For (2) and (3), one main challenge is the great diversity of access control models [16] proposed to describe policies. Policy-neutral access control (PNAC) languages [17] allow supporting several models, but lack real enforcement mechanisms. On the other end, several PNAC frameworks have been proposed [18], [19] but without generic enough specification languages. An interesting mid-term is described in [20] which combines a highly expressive security model (ABAC) [8], [21], a PNAC language (XACML) [22], and an authorization architecture. However, self-management of policies is not described. Further work is therefore needed.

III. Self-Protection Architecture

We now provide some background on the solution we explored for self-protection of pervasive networks [5]–[7].

We consider a pervasive system to be organized into a flat number of *clusters*, each containing a set of *nodes*. Nodes may join or leave a cluster dynamically. A *cluster* enforces a *cluster-level authorization policy*, applicable to nodes in the cluster. Nodes have various resource limitations, ranging from sensors to laptops, and enforce different *node-level authorization policies*. We now focus on a single cluster, but the approach can be generalized to any number of clusters.

For self-protection, we consider a security architecture divided into 3 abstract layers (see Figure 1). For each node, an *execution space* provides a running environment for application- or system-level services, encapsulated and manipulated as components. Node security management is performed in a *(security) control plane* using the VSK component. It oversees the execution space, both in terms of application-specific customizations and of enforcement of authorizations to access resources. Finally, a distributed *autonomic plane* supervises the VSK authorization policies in each node and performs the necessary adaptations at the cluster and node levels using several feedback loops. This paper provides an answer on how to design that layer using the ASPF framework.

VSK implements the managed OS-level security mechanisms in a node. It consists of a *Virtual Kernel (VK)* and an *Access Control Monitor* (ACM). The VK allows to reconfigure the *execution space* by providing run-time management functionalities over components and their bindings. It also efficiently controls access to the *execution space* resources, playing the role of an enforcement point for ACM decisions. Otherwise, the VK remains hidden in the background to minimize interactions with the *execution space* for performance optimization. The ACM is a decision engine allowing run-time selection of multiple authorization policies described in different access control models. Its design is compliant with the ABAC vision of access control [8], [21]: security attributes and permissions are separately managed (by an *Attribute Manager* for subject-attribute mappings, and by a *Rule Manager* for attribute-permissions assignments), and may be dynamically updated. More details regarding the VSK design may be found in [7].

We now present ASPF, a policy management framework which realizes an autonomic security manager above the VSK. The general idea is to adapt system security functionalities to the environment by context-aware change (tuning and/or generation) of authorization policies, such as adapting the policy strength to the ambient risk level.

The following sections describe the ASPF design principles, policy model (which specifies how to represent the managed authorization policies), security management framework, and resulting authorization architecture.

IV. ASPF Design Principles

We now describe the requirements for the framework.

A. Design Requirements

Several requirements should be met for such a self-protection framework.

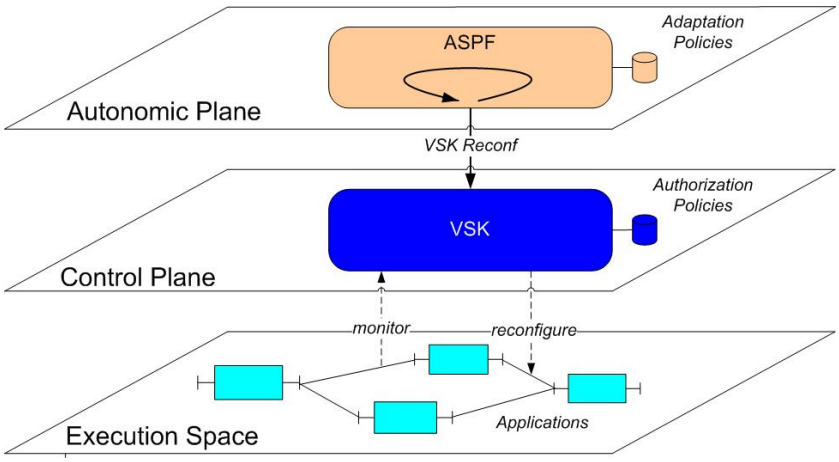


Figure 1. A 3-Level Architecture for Self-Protection.

1) Policy neutrality: Different types of authorization policies may be enforced in clusters and nodes. *Policy-neutrality* is thus mandatory to account for heterogeneous security domains by supporting several classes of authorization policies. Moreover, policies should be able to be *reconfigured dynamically* (including between different classes) when nodes move between security domains, or when the context changes.

2) Scalability: Pervasive systems are highly open and dynamic: nodes can enter and leave a network at run-time. The numbers of connected nodes may thus vary greatly in time, scaling network capacity both up and down, while the infrastructure remains unchanged. Scalability is thus a major challenge for the underlying protection framework, which should support both small- and large-scale systems.

3) Consistency: At the device level, a single system component (e.g., the security kernel [23]) usually controls all access to resources and enforces authorization policies. However, at the network level, each node still applies its own policy, but some nodes may share resources. The lack of a centralized module for enforcement of authorizations may lead to inconsistent network security policies. A solution for policy administration is thus required to guarantee consistency of distributed authorization policies.

4) User-friendly administration: Pervasive systems become increasingly complex, involving multiple users with different roles. Thus, the issue of system administration with minimal human intervention cannot be ignored. A security policy management framework should therefore simplify administration tasks and make system modifications transparent to users.

5) Context-awareness: Openness and dynamicity of pervasive networks induce rapid changes in the system context, calling for context-aware administration and protection. For instance, node availability may affect access privileges, as in ASRBAC authorization policies [24]. A node part of some clusters may have specific types of permissions that cannot be assigned to nodes in other clusters. Node migration between clusters may thus require update of access privileges. The management framework should thus select security functions based on evolution of the context.

6) Other Requirements: The security framework should also take into account requirements such as unified modeling of heterogeneous nodes, efficient protection mechanisms compatible with embedded constraints, or collaboration of decentralized security infrastructures.

B. ASPF Overview

Administration of authorization policies includes creation, deletion, and maintenance of access attributes and rules, and management of run-time constraints. To achieve this goal, ASPF applies the *autonomic approach* to make systems self-protected. Moreover, ASPF is *policy-driven*, i.e., the security behavior of the system is entirely governed by policies. The main distinguishing features of the framework are the following:

1) Policy-based management of authorization: The policy-driven approach is well adapted for administration of systems in open and dynamic environments: evolutions only trigger updates of applied policies, without changing the enforcement mechanisms. In our case, we use authorization policies to control protection. ASPF enables to modify, deploy, and enforce them through out the whole system.

2) Attribute-based authorization enforcement: Attribute-based access control [8], [20] is more suitable for open environments than traditional identity-based authorization: pervasive devices are not known by their exact names but by a dynamic set of *attributes*. This paradigm presents benefits in terms of expressivity and flexibility: it enables to support a large set of existing authorization policies, making policy-neutrality possible without developing a fully-fledged specific architecture. Separation of attributes from permissions also improves flexibility for dynamic policy reconfiguration.

3) Decentralized validation of authorizations: A scalable distributed system avoids using a central authority for validating authorizations. Our framework is based on a hybrid architecture using the concepts of *cluster* and *node*. Each node enforces a local authorization policy. Authorization policies of nodes inside a cluster are centrally controlled by a *cluster authority* which guarantees policy consistency between nodes. Policy synchronization between cluster authorities may be either centralized or decentralized. This architecture allows decentralized enforcement of authorization policies, while maintaining an efficient central control of policy deployment.

4) Integration of self-protection control loops: To satisfy the context-awareness requirement, ASPF regulates security using several self-protection feedback loops to select the authorization policy best fitting the system security context.

5) Self-configuration control loops for policy deployment: To guarantee consistency of decentralized policies, and facilitate system administration, self-configuration control loops allow the system to configure itself with minimal human intervention. Modification of chosen authorization policies will thus be automatically propagated through the whole network to guarantee consistent policy deployment.

V. ASPF Policy Model

In a pervasive system, different classes of authorization policies may be enforced: for instance, policies specified in the Domain and Type Enforcement (DTE) [25], Multiple Level Security (MLS) [26], or Role-Based Access Control (RBAC) [27] models.

ASPF allows to express those different models by describing authorization policies using the ABMAC (Attribute-Based Multi-Policy Access Control) model [20]. In this model, distinguishing features of system elements (subjects, objects, environment...) are described by attributes on which access decisions are based. Access attributes include principal identities, group membership, roles, security clearances, labels, or any other authorization information. Attributes are clearly separated from access rules, enabling independent modifications, e.g., activate/deactivate a role depending on location without reloading a full authorization policy.

```
<?xml version="1.0"?>
<DTEPolicy>
  <Rule>
    <Target>
      <SubjectAtt... name="domain">Trusted</SubjectAtt...>
      <ObjectAtt... name="type">Private</ObjectAtt...>
      <ActionAtt... name="operation">write</ActionAtt...>
    </Target>
    <Effect>grant</Effect>
  </Rule>
  ...
</DTEPolicy>
```

The ASPF policy model is shown in Figure 2. An ASPF policy consists of an *attribute map* and a set of *rules*. The map links system elements to their attributes. Elements may be subjects, objects, actions, or context data. Examples of

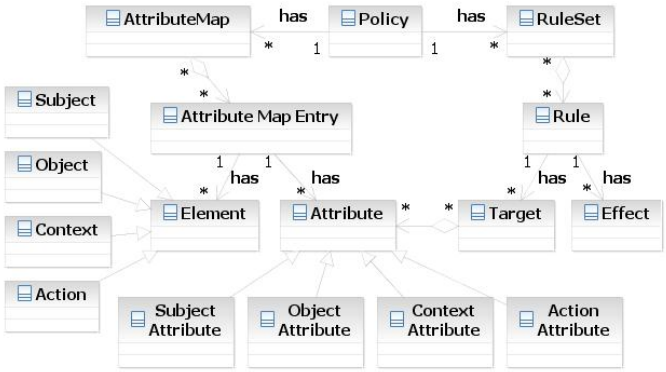


Figure 2. The ASPF Policy Model.

corresponding attributes include security domains, resource types, read/write operations, or location/time information. A rule contains a target, described by several attributes and an effect (allow/deny). A sample DTE policy is shown above, granting write authorizations to private resources for subjects in a trusted domain.

The result is a quite expressive model, while still remaining policy-neutral: as for XACML [22], specific authorization policies may be supported by refining the model through profiles. For instance, DTE, MLS, and RBAC policies are simply specified by defining the right types of attributes (domains, types, labels, roles...). Similarly, context-aware or history-based policies may be defined by adding specific context or history attributes.

As a drawback the processing of XACML policies may induce a performance overhead when dealing with policies with a large number of rules. This issue can be tackled by adopting an XACML policy optimization approach such as proposed in [28].

VI. ASPF Design

ASPF is a security management framework that governs authorization policies enforced by underlying VSK mechanisms. This section presents the ASPF design, based on several types of models.

A. Overall Design

The ASPF design is organized into three models:

- A *core model* describes system resources, security, and autonomic functionalities.
- An *extended model* refines the security and autonomic models for each type of resource.
- An *implementation model* describes the realization of the extended model, organizing functionalities into components to be implemented.

Those models are defined in the three steps shown in Figure 3. The core model consists of a *resource model*, a

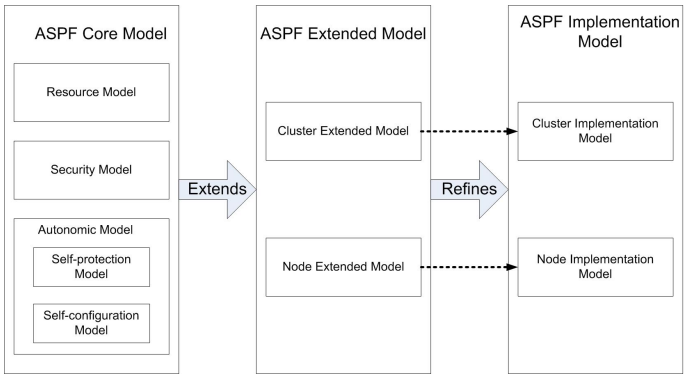


Figure 3. ASPF Overall Design.

security model and an *autonomic model*. These models are then refined into the *extended model* which involves a *cluster extended model* and a *node extended model* for cluster and node resources. Finally, these two models are refined into the corresponding implementation models.

B. ASPF Core Model

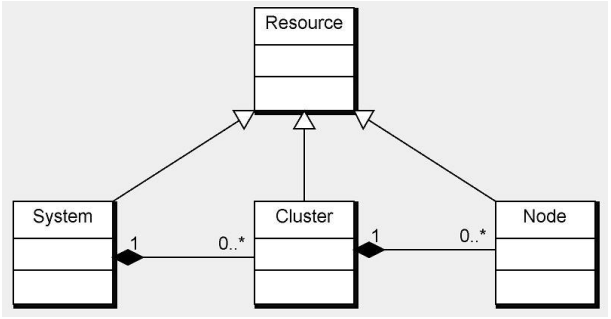


Figure 4. The Resource Model.

1) Resource Model: The *resource model* describes the structural organization of the system. The main concepts are those of *System*, *Cluster*, and *Node*, as shown in Figure 4:

- A *Resource* is the top-level concept which may be extended if the framework needs to be refined. It serves as coupling point with other models to describe different system functionalities.
- The *System* class represents the overall system to be protected (i.e., the pervasive network). It is organized into clusters.
- A *Cluster* is a coarse-grained structural unit including a set of nodes which collaborate to achieve some tasks, e.g., to provide a given service.
- A *Node* is the minimal structural unit. In pervasive networks, it represents a mobile device able to perform several functions and communicate with other nodes.

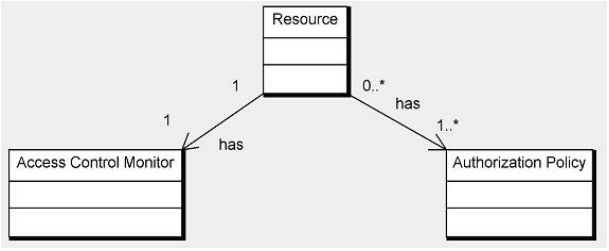


Figure 5. The Security Model.

2) Security Model: The *security model* specifies the authorization functionality to control access to *Resources*. The main concepts are those of *Access Control Monitor (ACM)* and *Authorization Policy* as shown in Figure 5:

- The *ACM* is a reference monitor which controls all access requests to resources.
- The *Authorization Policy* expresses conditions under which authorizations are granted or denied. It is specified according to the policy model previously described.

3) Autonomic Model: The *autonomic model* specifies how self-configuration and self-protection are achieved in the system. The *self-protection model* adapts authorization policies according to evolution of the context. The *self-configuration model* customizes authorization policies according to resource types, user preferences, or administrator-defined security governance strategies.

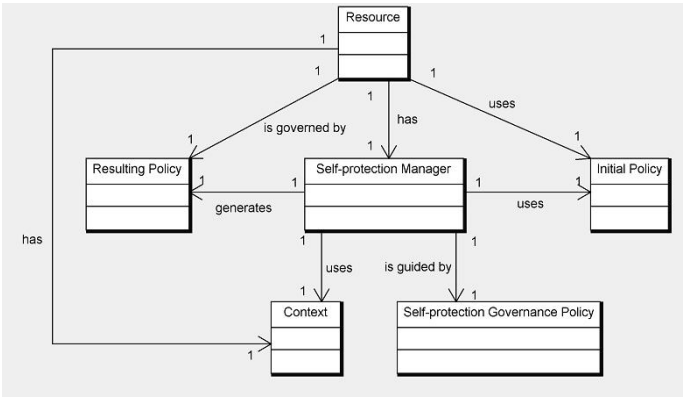


Figure 6. The Self-Protection Model.

The *self-protection model* describes how adaptations (selection of adequate security counter-measures) are launched at run-time, driven by evolution of the security context. Adaptations are performed both at the cluster level and at the node level. The main element of the model are the following, as shown in Figure 6:

- The *Self-Protection Manager* controls and orchestrates all activities related to self-protection. Its main role is to monitor the context and update authorization policies.
- The *Self-Protection Governance Policy* captures the administration strategy for self-protection. It drives

decision-making, specifying how to select the right authorization policy according to context information.
- The *Context* captures all information about the system environment which may influence such decisions.
- The *Initial Policy* is the current authorization policy, input for the context-aware security adaptation process.
- The *Resulting Policy* is the authorization policy output of the security adaptation process. This policy is generated by the *Self-Protection Manager*.

Once a *Resulting Policy* has been generated, this new policy should be propagated through the whole network for enforcement. As ASPF targets large-scale distributed systems, global policies should be translated into local ones to be enforced by each node. The *self-configuration model* specifies this translation process. The main element of the model are the following, as shown in Figure 7:

- The *Self-Configuration Manager* is the component in charge of the self-configuration process. It generates a *Resulting Policy* based on an *Initial Policy* according to a *Self-Configuration Governance Policy*.
- The *Self-Configuration Governance Policy* contains the guidelines for the translation process. It may be specified by condition-action rules.
- The *Initial Policy* is policy output of the self-protection model, input for the self-configuration process. It typically represents the new network security policy.
- The *Resulting Policy* is a policy derived from the *Initial Policy*, customized for each resource. It typically represents the new node security policy, adapted to the node-specific setting, e.g., by filtering all network access control rules not involving directly that node to comply with node computational limitations.

C. ASPF Extended Model

The role of the ASPF core model is to describe the security framework independently from the type of large-scale system. However, to be useful in practice, the framework must be described in a concrete setting. This is the purpose of the *extended model* which specifies the security framework for a specific type of large-scale system such as pervasive networking or cloud computing infrastructures. We now present an extended model for the pervasive setting which was the core focus of our study. However, another extension for cloud environments is detailed in Section X.

As pervasive systems are modeled as clusters and nodes, two extended models are defined to describe self-management of security at the cluster and node levels.

1) Cluster Extended Model: The main elements of the model are shown in Figure 8.

- The *Cluster Self-Protection Manager* captures the overall intelligence for self-protection of a cluster, coordinating the different necessary components.
- The *Cluster Context* class captures information about the context of a cluster. It may be specified using a more detailed context model such as DEN-ng [29] describing multiple dimensions of context.
- The *Cluster Self-Protection Governance Policy* captures the strategy to select the most adequate security function based on the cluster context.
- The *Cluster Initial Authorization Policy* is the starting point for the security adaptation process. It may be initially one of a set of predefined policies.
- The *Cluster Resulting Authorization Policy* is the result of the security adaptation process, and is generated by the *Cluster Self-Protection Manager* according to the current cluster status. That policy will then be applied to all nodes of the cluster.

The ASPF modular design into several models makes it more easy to select only the features necessary for the considered setting: compared to the core model, the cluster extended model only integrates the self-protection model. Authorization and self-configuration are left aside since: (1) policy enforcement is performed directly in the nodes; and (2) policy propagation towards nodes will be specified in the node extended model.

2) Node Extended Model: The main elements of the model are shown in Figure 9.

- The *Node Self-Configuration Manager* coordinates the components for self-configuration at the node level, i.e., to propagate adaptations decided at the cluster level. Such operations will be performed according to the *Node Self-Configuration Governance Policy*.
- The *Node Self-Protection Manager* orchestrates the components for self-protection of a node. Such operations will be performed according to the *Node Self-Protection Governance Policy* which describes reactions (i.e., authorization policies) to apply in security-sensitive situations, based on the *Node Context*.
- The *Node Resulting Authorization Policy* is the final output of the ASPF framework: after the adaptation process, both at the cluster and node levels, this policy will be installed inside the node for access control enforcement by the *Node Access Control Monitor*.

Overall, at the node level, self-management of security is a combination of self-configuration and self-protection: the result of the security adaptation process at the cluster level (*Cluster Resulting Authorization Policy*) is transformed into a *Node Resulting Authorization Policy* (self-configuration). Updates on the *Node Resulting Authorization Policy* will also be performed based on the node context (self-protection).

D. ASPF Implementation Model

The previous models are now refined at the implementation level, different implementation architectures being possible. In the sequel, we present an implementation model which fulfils the requirements presented in Section IV-A.

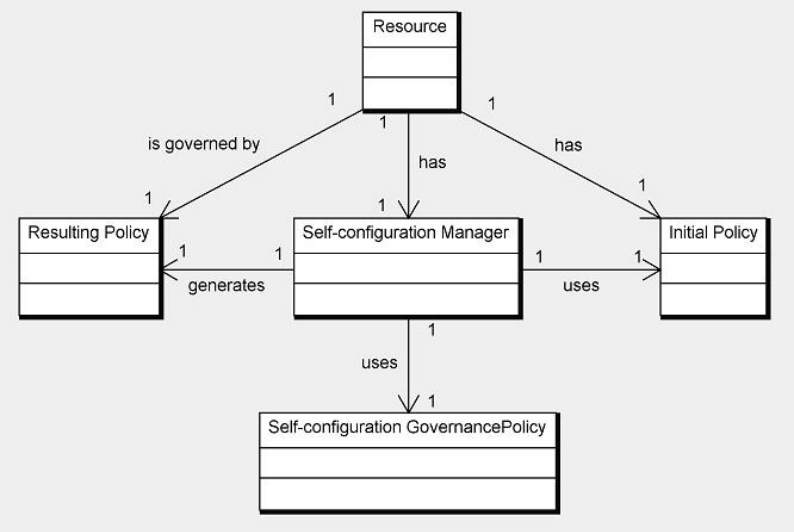


Figure 7. The Self-configuration Model.

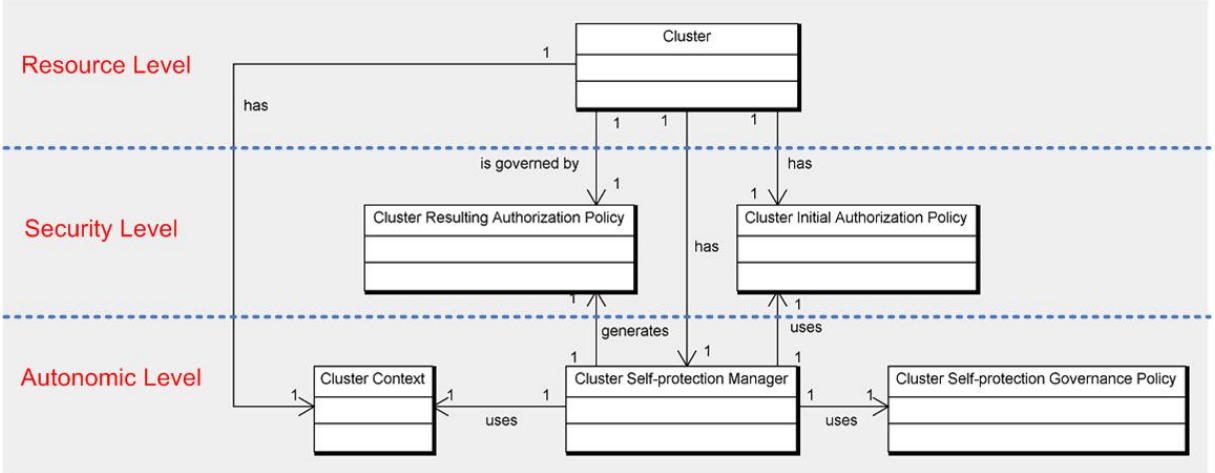


Figure 8. A Cluster Extended Model.

1) Cluster Implementation Model: The elements of the model are shown in Figure 10.

- The *Cluster Authority* component implements the *Cluster Self-protection Manager* class. It coordinates all self-protection tasks in the cluster.
- The *Cluster Context Monitor* provides a representation of the cluster security context. It aggregates low-level inputs from different sources (system/network monitoring probes, sensors,...), relying on context management infrastructures or intrusion detection systems.
- The *Cluster Authorization Policy Repository* contains a set of initial cluster authorization policies to enforce protection within different potential situations.
- The *Cluster Governance Policy Engine* generates security adaptation strategies to tune authorization policies to the environment, e.g., use DTE (resp. MLS) policies in a friendly (resp. hostile) setting. It may also define new policies to cope with unknown situations.
- The *Cluster Resulting Authorization Policy* is the output of the cluster-level security adaptation process.

2) Node Implementation Model: The main elements of the model are shown in Figure 11.

- The *Self-Configuration Manager* and *Self-Protection Manager* functionalities are implemented by two components, the *Node Authority* and the *Node Adapter*. The *Node Authority* typically resides on a server at the cluster-level, while the *Node Adapter* is a component local to each node. The *Node Authority* is the main control point to administer security configurations and customize authorization policies. The *Node Adapter* is a local security controller in the node with two roles. It is a proxy for the remote *Node Authority*, executing its

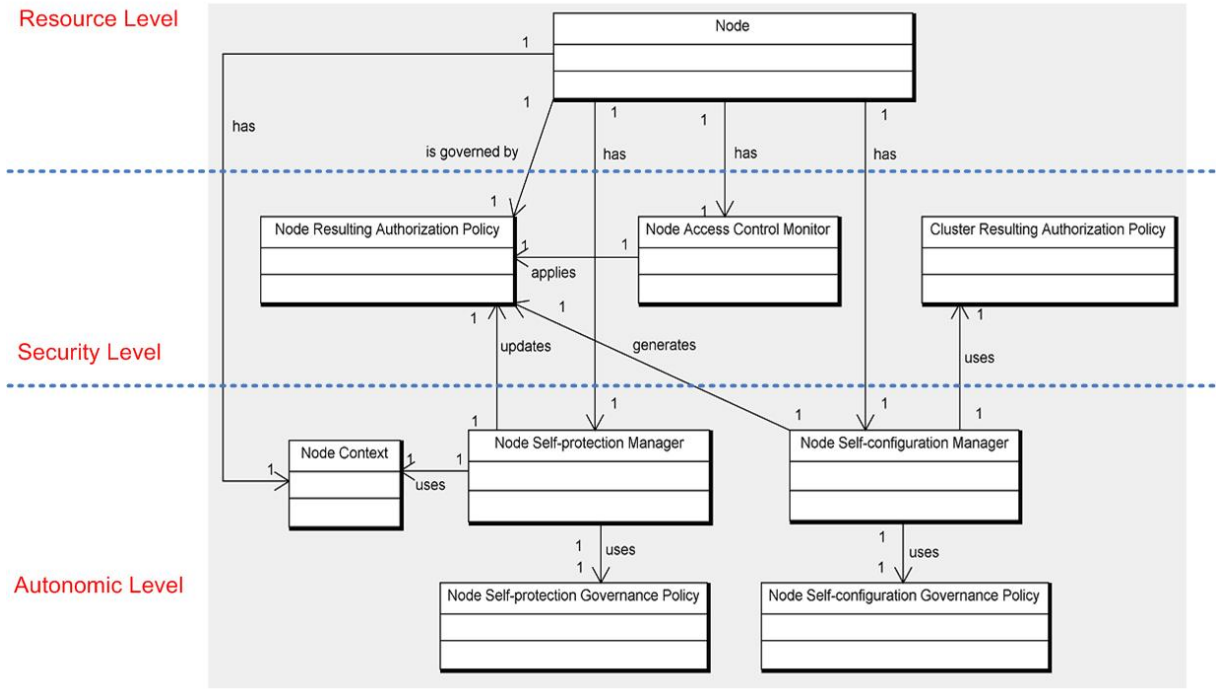


Figure 9. The Node Extended Model.

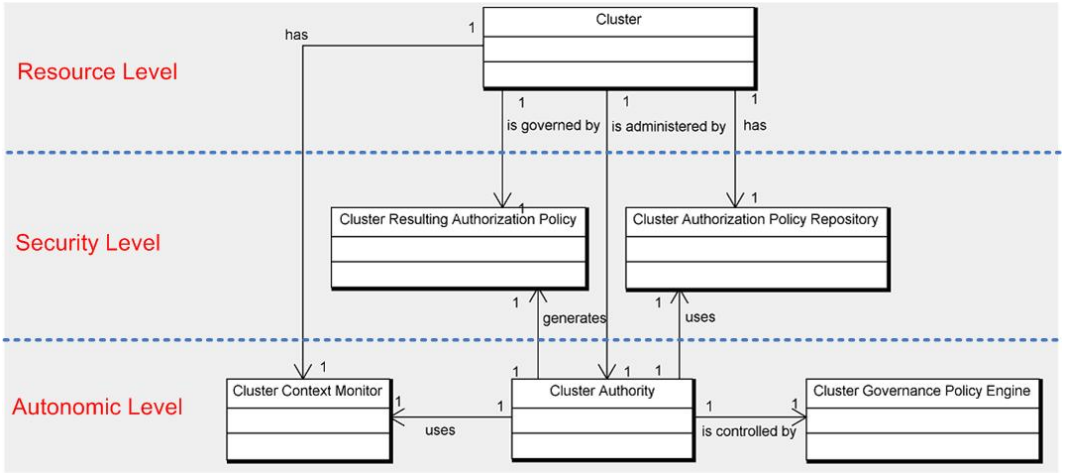


Figure 10. The Cluster Implementation Model.

decisions and installing in the node authorization policies customized at the other endpoint. It also realizes node-level self-protection to adapt node authorization policies based on the node context.

- The *Node Profile* refines the *Node Self-Configuration Governance Policy* by describing the node capabilities (CPU, memory, storage...). As a cluster might contain many nodes, a large part of cluster policy rules might not be relevant for each node and should be filtered. In our design, node-level self-configuration is viewed as filtering the cluster authorization policy according to constraints described in this profile.
- Other components such as the *Node Context Monitor* or the *Node Governance Policy Engine* play the same roles as on the cluster side, but for the node setting.
- The *Node Resulting Authorization Policy* is the final output of the node-level security adaptation process. The corresponding access control rules may then be enforced in the node with a lightweight authorization overhead thanks to the underlying VSK OS.

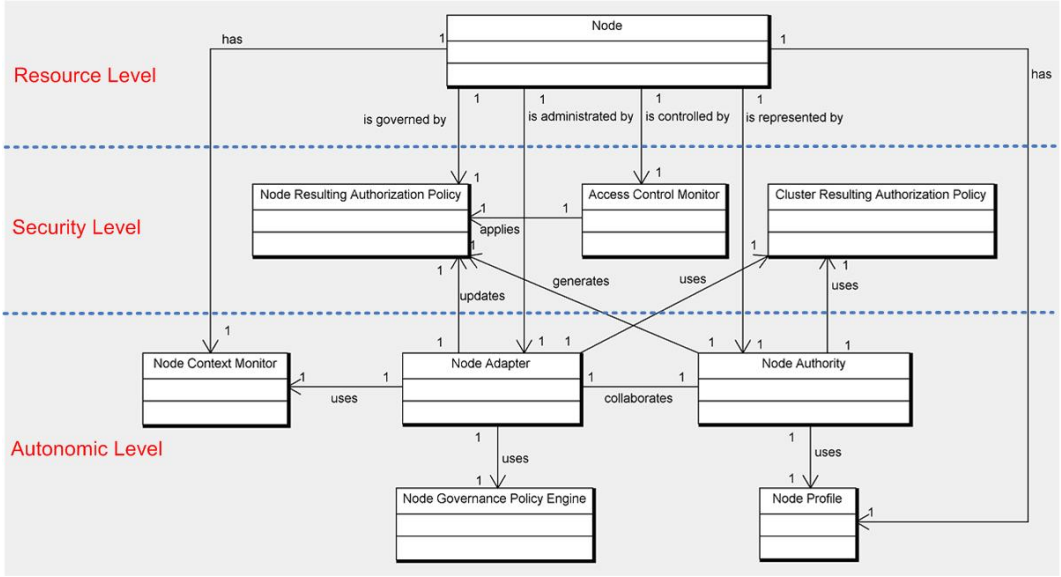


Figure 11. The Node Implementation Model.

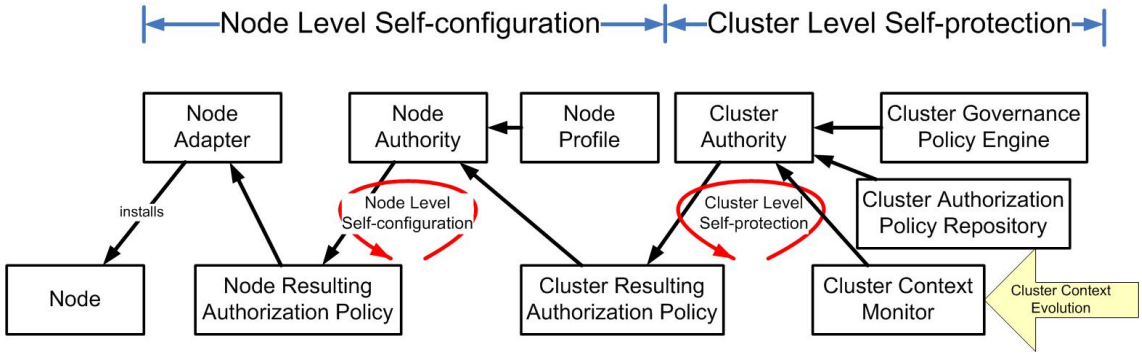


Figure 12. Cluster-Level Self-Protection Loop

E. A Double Control Loop for Self-Protection

ASPF regulates security at two levels, using separate feedback loops, both at the cluster and node levels. The previous implementation components interact as follows.

1) Cluster-Level Self-Protection: This loop shown in Figure 12 aims to mitigate threats against a cluster. The *Cluster Context Monitor* aggregates security-relevant events to reach a representation of the cluster security context. It notifies the *Cluster Authority* in case the context changes. The *Cluster Authority* then updates the *Cluster Authorization Policy*, according to the strategy specified in the *Cluster Self-Protection Governance Policy*. This operation may be performed by selecting a predefined stronger/weaker policy from the *Cluster Authorization Policy Repository*.

Nodes have severe resource limitations, for instance in terms of computing capabilities or power consumption. Execution must therefore be optimized. The chosen cluster policy is further interpreted by each node according to its specificities (CPU, memory, battery, etc.) captured in the *Node Profile*, generating a new node authorization policy (*Node Resulting Authorization Policy*). Policy rules not relevant for each node are notably discarded. This policy is then loaded into the node authorization sub-system for enforcement. This customization improves efficiency and scalability. It also makes the system more manageable by reducing the number of authorization rules.

2) Node-Level Self-protection: A simpler loop also operates at the node level to defeat attacks on a single node as shown in Figure 13). Based on information about the node security context (captured by the *Node Context Monitor*), this loop tunes security attributes such as assigning a different role to a subject in order to reduce his privileges in a hostile environment, without modifying the rest of the node authorization policy. For instance, when a node is attacked, the security level of a highly sensitive resource could be increased from *Confidential* to *Top Secret* to minimize possibilities to access the resource.

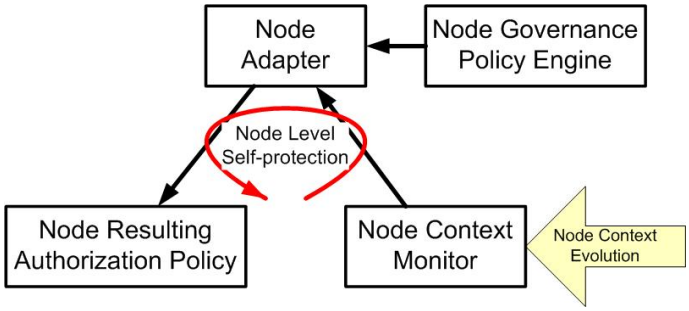


Figure 13. Node-Level Self-Protection Loop

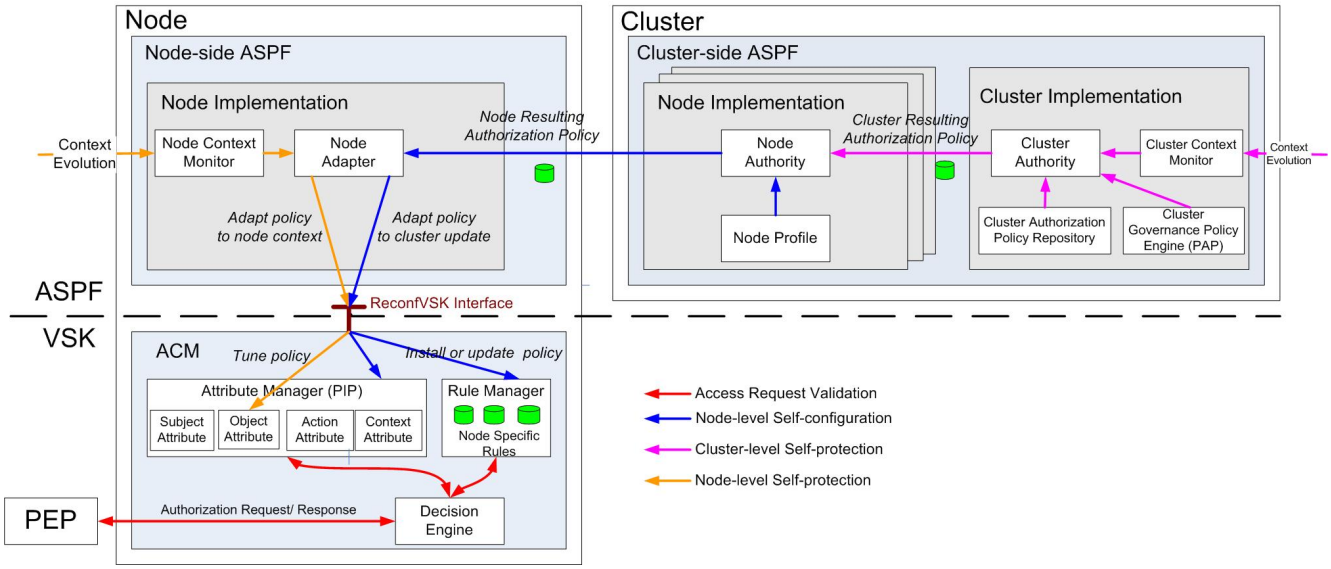


Figure 14. Authorization Architecture

VII. Authorization Architecture

The ASPF authorization architecture integrates the self-configuration and self-protection models into the XACML authorization framework (see Figure 14). XACML defines 4 main components for policy enforcement (PEP), decision-making (PDP), administration (PAP), and management of attributes (PIP) [22]. In the VSK architecture: the PEP is the *Virtual Kernel (VK)* component, which enforces authorizations on execution resources; the PDP is the *Decision Engine* component; the PIP is the *Attribute Manager (AM)* component that provides additional information for access validation. The authorization policy is stored in the *Rule Manager (RM)* component.

Access requests to resources (located in the execution space) are forwarded to the Decision Engine and transformed to an ABAC-compliant request. Attributes are fetched from the Attribute Manager, and the request is validated against the authorization policy. The decision is then enforced by the VK by reconfiguring the execution space to establish access to requested resources.

ASPF may be seen as an enhanced PDP. Pure decision-making is extended with autonomic capabilities to generate or tune the security policies contained in the VSK ACM based on policy sets written by a cluster network administrator.

Figure 14 shows how ASPF realizes the two self-protection control loops described in Section VI-E. The cluster-level self-protection model together with the node-level self-configuration model achieve a global control loop which updates both rules and attributes of authorization policies according to cluster context and node profile information. The node-level self-protection loop tunes security attributes based on node context information. The overall architecture not only performs access control enforcement and decision-making. It also improves management of authorization policies, notably by enabling context-aware adaptations thanks to autonomic features.

VIII. Implementation

A first prototype implementation was realized including a set of devices running a kernel composed of VSK and of a node-side ASPF component (implemented using the THINK [30] component-based OS framework), and a cluster authority server (implemented in Java). DTE, MLS and RBAC security policies are currently supported, with 10 subjects (threads) and 60 objects (system calls) to model a typical embedded OS environment, and 3 security levels for the cluster security context. Cluster policies are filtered according to active subjects or objects described in the node profile. The resulting attribute mappings and rules are then loaded inside the VSK via a dedicated reconfiguration interface `ReconfVSK` allowing to change dynamically security attributes and policy rules.

IX. Evaluation

The self-protection capabilities of the framework (ASPF+VSK) were evaluated in terms of overall response time and resiliency to attacks. A qualitative assessment of the security of the framework is also given. All measurements were performed on a 2.7GHz DELL OptiPlex 740 desktop PC with Linux/Ubuntu 9.04 and 1GB of RAM, on which are run the cluster authority server and node simulations.

A. End-to-End Response Time

We measure the overall latency to complete a full self-protection loop for adaptations at the cluster and node levels. Evaluation results for each step of the loop are shown in Figures 15a and 15b for different types of security policies.

In the first benchmark, detection of an attack on a cluster of 100 nodes in a steady state is simulated by direct update of the cluster security context. In practice, this step would be performed by an Intrusion Detection System (IDS) such as Snort, with 1ms as typical order of magnitude for attack detection and countermeasure initiation. The next steps are generation of a node-specific policy (given times are averaged on the number of nodes), invoking the node VSK to load the policy, kernel reconfiguration with the new policy, and return to the steady state. The overall latency averaged over different security models is 33.92ms.

In the second benchmark, attacks are detected by a node context handler. The next steps include invoking the VSK, tuning security attributes to adapt to the new security context, and returning to a steady state. The measured overall latency for this adaptation loop is 1.15ms.

Overall, the adaptation response times seem reasonable, since the time between two policy reconfigurations is typically from a few seconds to one minute, for instance when switching between wireless networks in different locations. As expected, node-level adaptations are much lighter than cluster-level reconfigurations. This is in part due to the ABAC approach: the same authorization rules may be applied, only attributes values being tuned. For highly dynamic environments, this design makes self-protection more lightweight, allowing to follow small variations of the context, without regenerating a full policy.

B. Resilience

To measure the effectiveness of self-protection using ASPF, we use the methodology for benchmarking self-* capabilities of autonomic systems proposed in [31] based on injection of disturbances (see Figure 16a). The idea, coming from dependability benchmarks, is to introduce in the System Under Test (SUT) disturbances in the form of attacks or faults, and to measure the impact on the performance workload. This type of benchmark, already used to assess self-healing abilities, measures how well the SUT adapts to the injected changes in terms of speed of recovery, impact on performance, etc.

In our case, the SUT is the set VSK+ASPF on which is applied a workload to validate access requests from the execution space. We measure the impact on throughput (number of requests per second validated by the VSK, averaged over a sliding sampling time window τ) of updating security policies to respond to injected attacks. An attack from a malicious node is simulated by directly changing the cluster security context at the beginning of an injection slot, and waiting from the SUT to come back to a steady state. The results are shown in Figure 16b for $\tau = 1ms$ and $\tau = 0.16ms$, which is about the latency value for an end-to-end reconfiguration. The decrease in throughput due to security adaptations depends on the sampling slot value: 89% for $\tau = 0.16ms$ (worst case), but only 15% for $\tau = 1ms$ (standard situation). These results show that the system is able to protect itself effectively with a reasonable performance cost. The recovery time is almost immediate for $\tau = 0.16ms$, and about 2ms for $\tau = 1ms$. Thus, the system is able to complete successfully its reconfiguration in times which are largely acceptable. These metrics tend to show that ASPF provides self-protection with minimal impact on system resources.

C. Security Evaluation

Evaluating the quality of the autonomic response is harder: does the system remain secure after a security reconfiguration? To avoid rogue third parties to directly update node authorization policies inside the VSK, a single reconfiguration interface (`ReconfVSK`) is introduced as unique entry point to control the VSK. This interface remains internal to a node, to avoid policy update requests coming from the network aiming to lower node security settings. ASPF behaves as a distributed security management plane which guarantees complete mediation over this interface: all authorization policy modifications may only be issued by the *Node Adapter*, *Node Authority*, and *Cluster Authority*

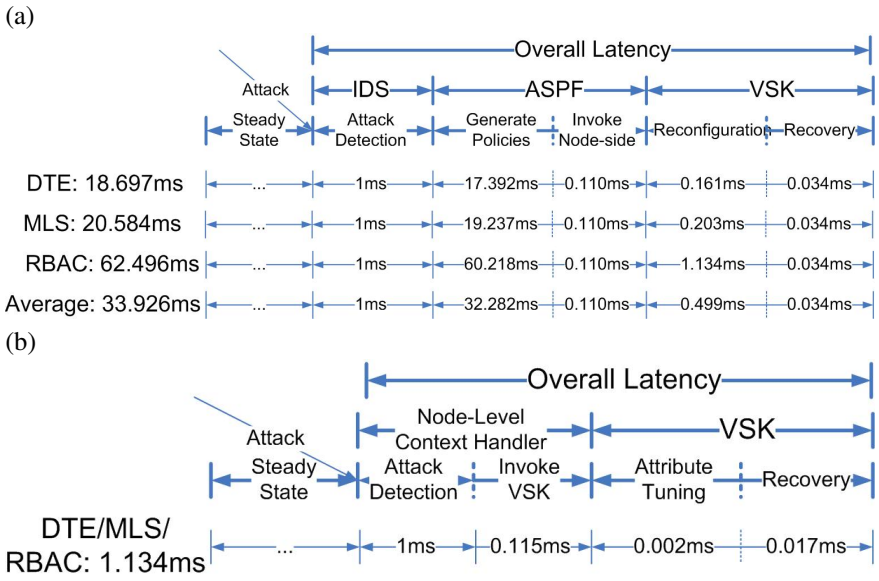


Figure 15. Self-Protection Latencies: (a) Cluster-Level; (b) Node-Level.

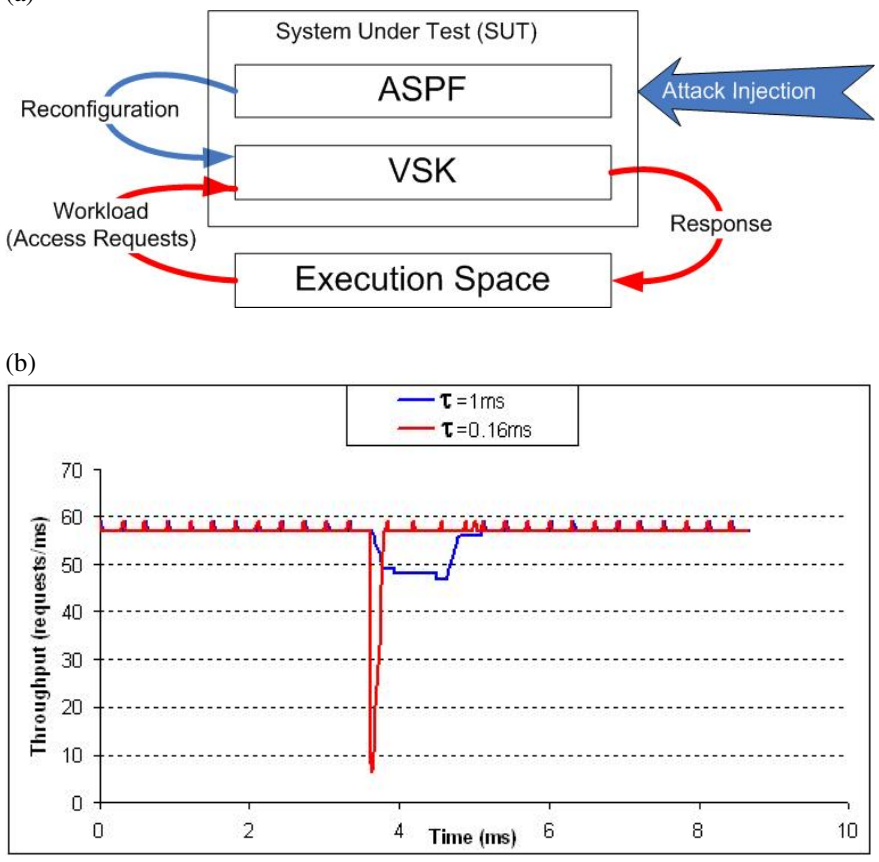


Figure 16. Benchmarking Self-Protection Capabilities: (a) Principle; (b) Results.

components along a trusted path. The link between node-side and cluster-side ASPF components is also assumed to be a secure, authenticated channel to avoid man-in-the-middle attacks or rogue cluster authorities. Finally, an MMU-like hardware mechanism in the node prevents circumventing the *Node Adapter* component. These features qualify ASPF as a strongly protected management plane over VSK authorization mechanisms.

The underlying VSK infrastructure which serves as foundation for ASPF was also evaluated in terms of security. Three main threats were identified:

1) An application-level component gains illegitimate access rights through an existing binding. For bindings, the current THINK framework does not distinguish between *read* and *write* actions, i.e., a binding for a *read* access action could be used to perform a *write* invocation. Thus, a malicious component could perform a privilege escalation through such a binding, as there is currently no checking mechanism in THINK to prevent it.
2) An application-level component illegally accesses another such component by bypassing VSK protection. The VSK limits itself to checking and validating access requests based on presented attributes. However, an illegal access bypassing the kernel remains possible, as in THINK an invocation may directly access to a physical address without any control.
3) An application-level component illegally accesses the VSK. This threat is an extension of the previous one: since VSK is also built on the THINK framework, access to the kernel by directly jumping to a physical address may be possible.

The first threat may be mitigated by extending the definition of interfaces of the THINK framework with access action types. During compilation, checking may be included to determine if method invocations match authorized access types. The second and third threats correspond to bypass attacks. A MMU-based hardware mechanism is usually used to avoid circumventing reference monitors. Such mechanisms may be used to prevent bypass of VSK authorization checks. One MMU solution for component-based OSes was implemented in *CRACKER* [19]: the MMU organizes components into different memory pages according to their security level, and performs additional checking for inter-page invocations. For some hardware platforms like AVR or ARM which do not support MMU, a tool was proposed for code checking which replaces memory access by a pointer to a manager for security policy validation [32]. We believe that isolation between application-level components and the VSK may be achieved through these two categories of solutions.

X. Applying ASPF To Cloud Infrastructures

We now further validate the framework design by showing through a short case study that ASPF is generic enough to be applicable to other types of large-scale systems than simply pervasive networks. In the sequel, we focus on cloud computing infrastructures. We first recall some of the main security issues of those environments (Section X-A), highlighting the need for self-protection mechanisms. We then present the targetted self-protection scenarios (Section X-B). We finally show how the ASPF core model (Section X-C), extended model (Section X-D), and authorization architecture (Section X-E) may be refined to realize and coordinate several self-protection loops in a cloud setting.

A. Towards Self-Protecting Clouds

Cloud computing raises many security challenges [33], notably due to vulnerabilities introduced by virtualization of computing resources, and unclear effectiveness of traditional security architectures in fully virtualized networks. One of the main issues is how to guarantee strong resource isolation, both on the computing and networking side in a multi-tenant environment.

Few solutions are available, usually addressing only one of the two aspects [34], [35]. The extremely short response times required to activate system defenses efficiently, and the impossibility of manual security maintenance call for a flexible, dynamic, and automated security management of cloud infrastructures, which is clearly lacking today. A framework enabling self-protection of a cloud infrastructure could provide answers to some of those challenges, making ASPF an interesting candidate to reach this objective.

In the cloud, virtualization has two facets:

- *Computing resources* are abstracted away from the hardware in the form of *virtual machines (VMs)* isolated by a hypervisor on each server of a data center. Threats come at two levels of granularity: at the host level, through weaknesses either in the VM (guest OS) or the hypervisor; and at the cloud-level, mainly in the form of network-level attacks found in traditional security environments (e.g., DDoS). An autonomous security management framework for the cloud should thus put in place self-protection loops at each of those two levels.
- *Network resources* (routers, firewalls,...) themselves become virtualized, e.g., as virtual appliances. Network zones where traffic could be separated physically or logically using VLANs or VPNs are replaced by *logical security domains* which may have variable boundaries. It is thus critical to be able to manage security autonomously in such "islands". The security management framework should thus also provide self-protection abilities in logical security domains, called *VSBs (Virtual Security Domains)* in the sequel.

B. Cloud Self-Protection Scenario

We explore the realization of *adaptable quarantine zones*: a number of VMs considered as compromised are isolated from the data center temporarily. Confinement may be lifted when the risk has decreased, and the VMs not considered hostile any more.

We assume that on each physical machine of the data center is installed a firewall component which allows to control strictly communications between VMs: an authorization policy specifies which interactions are allowed/forbidden. This virtual firewall may for instance be located in the

domain 0 of a Xen hypervisor. Additional firewalls may also be placed at the cloud level to control inter-machine communications. The authorization policy is reconfigurable dynamically according to the estimated level of risk. Self-protection of the virtualized infrastructure then consists in adapting this set of policies according to the execution context of the data center, more or less hostile. Depending on alerts generated from an IDS (local or distributed in the data center), the most adequate authorization policy is autonomously selected, and installed in the different firewalls to realize hardened control over VM communications, and enforce the quarantine zone (see Figure 17).

In what follows, the quarantine zone is implemented at three levels of granularity: (1) within in physical server (*machine-level self-protection*); (2) within a VSB (*logical self-protection*); and (3) at the cloud level (*system-level self-protection*). The next sections describe how the ASPF core and extended models may be refined to realize those 3 self-protection loops.

C. ASPF Core Model

1) Resource Model: This model describes the organization of a cloud infrastructure (see Figure 18). As for the pervasive case, entities derive from a generic *Resource* class.

- The *System* class represents the overall cloud infrastructure to be protected, physically composed of a set of machines and logically divided into several *VSBs*. Both physical and logical isolation are realized through *Authorization Policies*.
- A *Machine* is a server in the data center. It hosts several *Virtual Machines (VMs)*, isolated by an hypervisor, which may create, destroy, or migrate VMs on demand.
- The VM is the first-class architectural component of the cloud. It runs a guest OS on top of the hypervisor, which manages VM resources.
- The VSB is a logical unit of VM isolation, e.g. to compartimentalize different services. VMs belonging to a VSB may be distributed on several machines. VSBs may be strictly isolated between each other using network-level mechanisms.
- A *Local VSB* contains all VMs of a VSB which reside on a given machine. It realizes local isolation from VMs of other VSBs in the machine. VM isolation at the VSB level is achieved by collaboration between all the corresponding Local VSBs.

2) Security Model: As for the pervasive case, access to resources is controlled by authorization policies. However, in the cloud, the security model features several types of policies since the resource model is richer (see Figure 19).

- The *System Authorization Policy* contains all access permissions to cloud resources. It will be enforced by the *System ACM* component at the cloud level.
- The *VSB Authorization Policy* contains access permissions in the scope of a VSB: it controls VM access at a logical level (the VSB security domain), regardless of the VM physical location. If we assume that access between two VMs belonging to different VSBs is always denied (strict isolation between VSBs), the System Authorization Policy may be viewed as the collection of VSB Authorization Policies. Policies in each VSB may be specified in different authorization models (e.g., DTE, MLS, or RBAC), as each VSB is a security island where policies may be administrated in a specific manner.
- The *Local VSB Authorization Policy* is the projection of the VSB Authorization Policy inside a machine, and thus corresponds to two types of situations: VMs are co-located on the same machine; or VMs reside in different machines. In the former situation, access may be directly validated by at the machine-level. The latter calls for inter-machine collaboration.
- The *Machine Authorization Policy* is the collection of Local VSB Authorization Policies for all Local VSBs in the machine. Due to possible heterogeneity of authorization models between VSBs, in the general case, the Machine Authorization Policy will be a set of Local VSB Authorization Policies specified in different models. This policy will be enforced by the *Machine ACM* component residing on each machine.

In our cloud model, to control inter-VM communications, policy enforcement is performed both at the machine level and the system level. We describe next a simple solution, other alternatives being possible.

If the VMs reside on the same machine, the Machine ACM applies the Machine Authorization Policy to validate the request. Since by default the VMs reside in the same VSB, validation is straightforward by enforcing the corresponding Local VSB Authorization Policy. However, since Local VSB Authorization Policies may be described in different models, a policy-neutral solution is required for access control enforcement at the machine level. Using ABAC for policy specification allows to achieve that goal as in the pervasive case.

If the VMs reside in different machines, the Machine ACM of the requesting VM checks in its Machine Authorization Policy whether this VM has permission to access an external machine. Control is then transferred to the System ACM which checks in the System Authorization Policy whether inter-machine communication to the target VM is allowed. Finally, the Machine ACM of the target VM checks that requests to this VM coming from a remote machine are allowed. Such a three-step validation of requests allows authorization to be more efficient and scalable (local policies do not deal with inter-machine communications) and to check consistency of distributed policies at the system level.

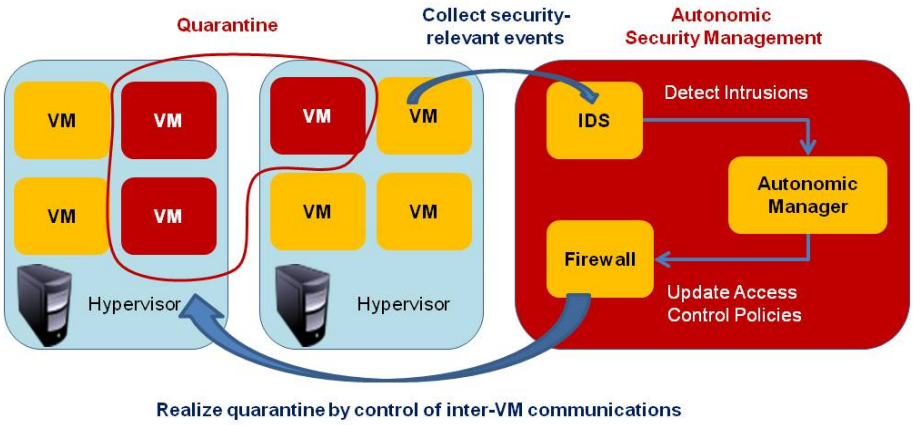


Figure 17. An Adaptable Quarantine Zone.

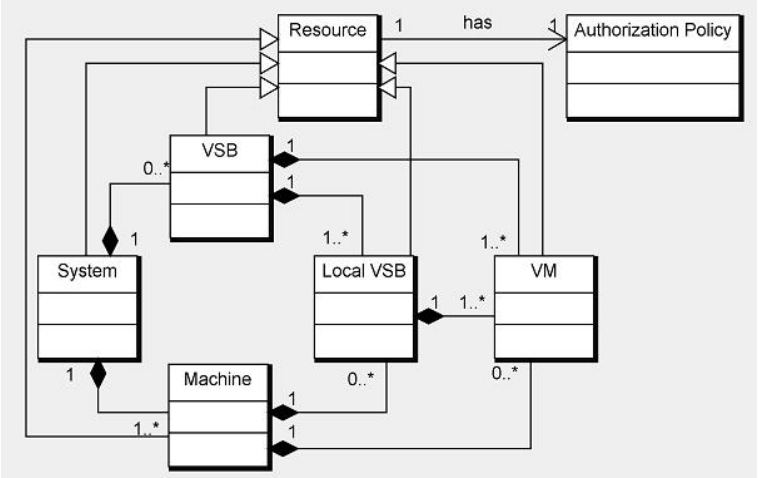


Figure 18. Cloud Resource Model.

D. Extended Models

The extended models describe the realization of several self-protection loops at different levels of granularity in the cloud, to address threats targeted at a machine, a logical security domain (i.e., a VSB), or the cloud itself by updating the corresponding authorization policies.

1) Machine Extended Model: If a malicious VM compromises the hypervisor [36], [37], the threat may spread to all the VMs residing on the machine, which may need to be confined. Defeating such attacks is the objective of this self-protection loop (Figure 20).

When an attack is detected by the *Machine Context* monitor, the *Machine Self-Protection Manager* applies a *Machine Self-Protection Governance Policy* to adapt the Machine Authorization Policy to the current situation, policy which will be propagated to the authorization policies of each Local VSB on the machine. At the same time, the manager collaborates with the *System Self-Protection Manager* to determine whether further counter-measures should be triggered at the cloud level.

2) VSB Extended Model: This self-protection loop (Figure 21) addresses a wider scope: it aims to defeat attacks which have spread into a logical security domain, e.g., by isolating compromised VMs. The VSB Authorization Policy is updated to fit the evolving *VSB Security Context* – those modifications are propagated to the System Authorization Policy to maintain policy consistency. A self-configuration loop is then launched to refine this policy into corresponding Local VSB Authorization Policies – the modifications being propagated to the Machine Authorization Policies.

3) System Extended Model: Two events may launch the system self-protection loop (Figure 22): detection of a cloud-level attack through *System Context* monitoring; or a request from a Machine Self-Protection Manager for increased counter-measures, faced with an anomaly which cannot be handled at the machine level alone. Regarding self-protection, the *System Self-Protection Manager* tunes the System Authorization Policy following the run-time

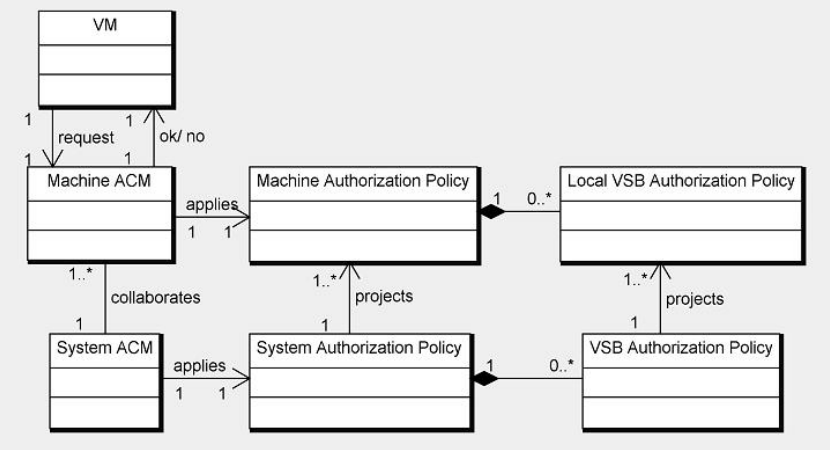


Figure 19. Cloud Security Model.

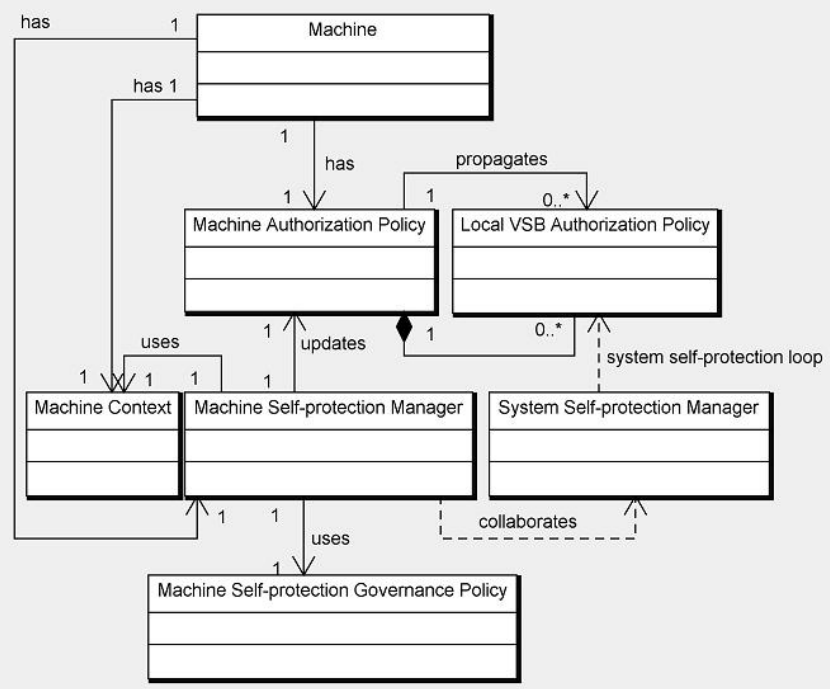


Figure 20. Machine Extended Model.

adaptation strategy defined in the *System Self-Protection Governance Policy*. This update is propagated towards the relevant VSB Authorization Policies. As in a pervasive case, on each machine, a self-configuration mechanism then translates each VSB Authorization Policy into a Local VSB Authorization Policy, finally updating the Machine Authorization Policy.

E. Authorization Architecture

An authorization architecture called SECloud was defined to implement the previous self-protection models. SECloud refines the ASPF authorization architecture. As shown in Figure 23, authorization validation is the result of a collaboration between System and Machine ACMs. SECloud consists of a number of server-side components installed in the cloud service provider network to control System, VSB, and local VSB functionalities, while the machine-side components essentially apply authorization policy adaptation decisions taken at the other end-point, and control access among local VMs. Such an architecture is currently under implementation.

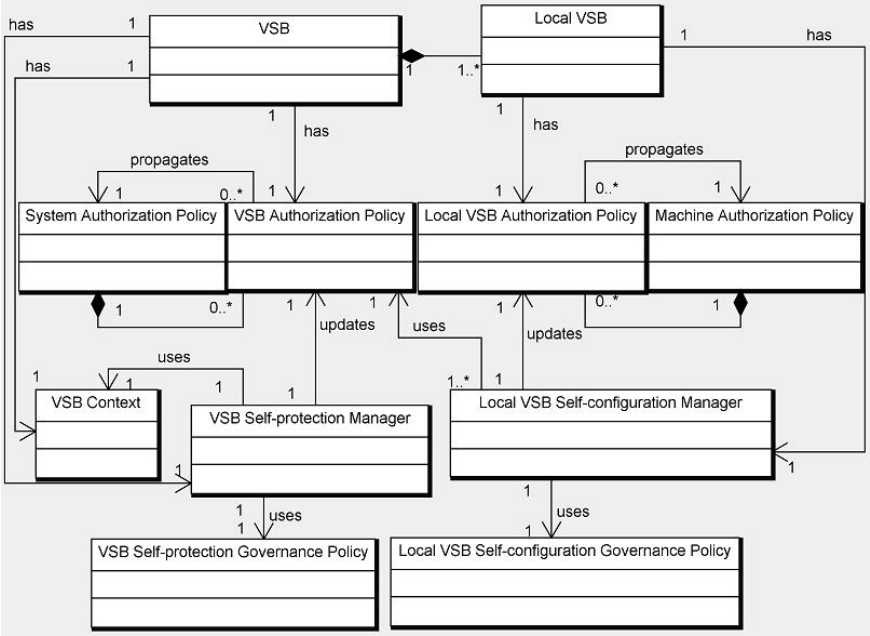


Figure 21. VSB Extended Model.

XI. Conclusion

This paper presented ASPF, a policy-based security management framework illustrating the design of an autonomic security manager to control OS-level authorization mechanisms in a pervasive device. ASPF implements several self-protection loops, authorization policies being adapted according to security context variation at both the network and device levels. Policies are described with an attribute-based extension of XACML to support policies specified in multiple authorization models. Performance, resilience, and security evaluations show that, together with VSK, ASPF provides strong and yet tuneable security while still achieving good performance, making it suitable for self-protection of pervasive systems. ASPF is also applicable to other types of large-scale systems such as cloud computing environments.

Current work focuses on the definition of the security adaptation strategy. We are currently investigating the approach where autonomic management strategies are specified using domain-specific languages (DSLs) [38]. Current ASPF adaptation strategies are purely action-based. However, higher-level strategies using objective or utility function policies are also desirable [39]. By enabling the specification of governance strategies with richer types of policies, the DSL approach should allow describing self-managed security at different levels of granularity which can be refined (e.g., with notions of policy continuum [40]), and thus evolve towards greater autonomic maturity in the corresponding systems.

Acknowledgments

This work has been funded by the ANR SelfXL project.

References

[1] R. He, M. Lacoste, and J. Leneutre, “A Policy Management Framework for Self-Protection of Pervasive Systems,” in *International Conference on Autonomic and Autonomous Systems (ICAS)*, 2010.

[2] D. Chess, C. Palmer, and S. White, “Security in an Autonomic Computing Environment,” *IBM Systems Journal*, vol. 42, no. 1, pp. 107–118, 2003.

[3] IBM, “An Architectural Blueprint for Autonomic Computing,” 2006, Autonomic Computing White Paper.

[4] J. Strassner, *Policy-Based Network Management: Solutions for the Next Generation*. Morgan Kaufman, 2003.

[5] R. He and M. Lacoste, “Applying Component-Based Design to Self-Protection of Ubiquitous Systems,” in *3rd ACM workshop on Software Engineering for Pervasive Services (SEPS)*, 2008.

[6] R. He, M. Lacoste, and J. Leneutre, “An OS Architecture for Device Self-Protection,” in *International Symposium on Stabilization, Safety, and Security of Distributed Systems (SSS)*, 2009.

[7] ——, “Virtual Security Kernel: A Component-Based OS Architecture for Self-Protection,” in *3rd IEEE International Symposium on Trust, Security and Privacy for Emerging Applications (TSP)*, 2010.

[8] L. Wang, D. Wijesekera, and S. Jajodia, “A Logic-Based Framework for Attribute-Based Access Control,” in *ACM Workshop on Formal Methods in Security Engineering*, 2004.

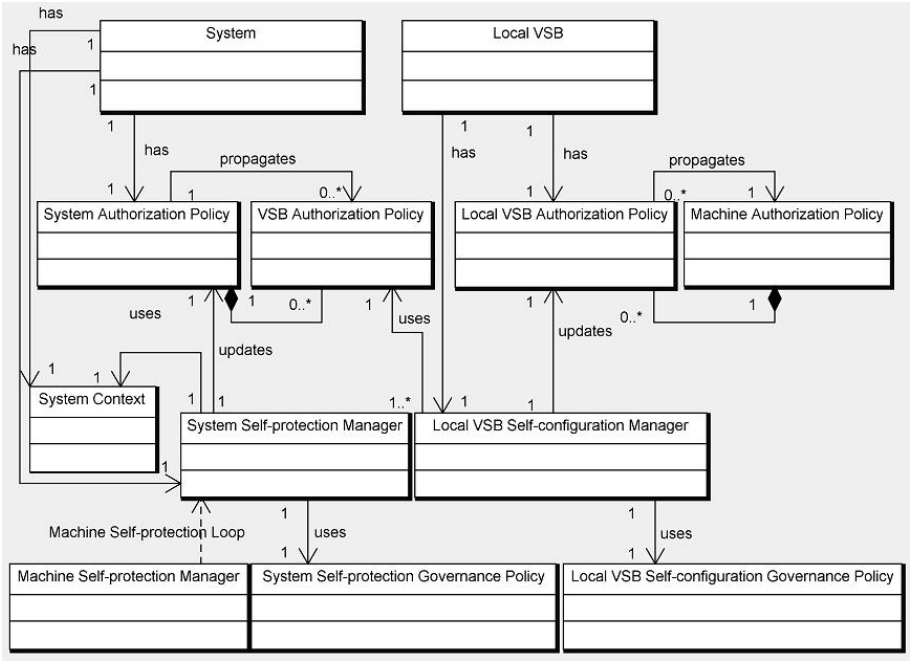


Figure 22. System Extended Model.

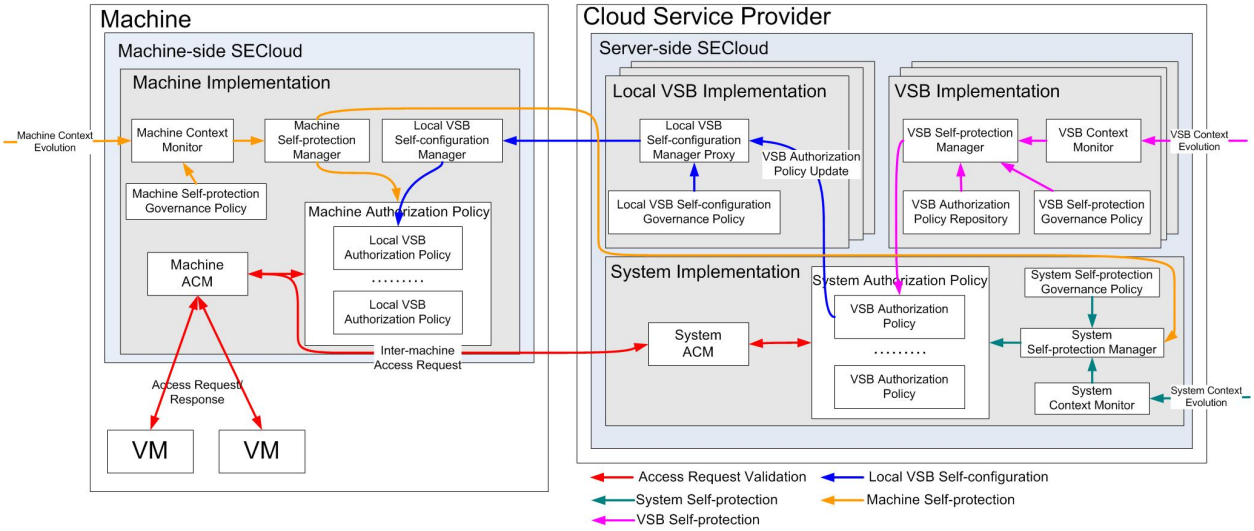


Figure 23. The SECloud Authorization Architecture.

[9] B. Claudel, N. De Palma, R. Lachaize, and D. Hagimont, "Self-Protection for Distributed Component-Based Applications," in *International Symposium on Stabilization, Safety, and Security of Distributed Systems (SSS)*, 2007.

[10] J. Agosta et al., "Towards Autonomic Enterprise Security: Self-Defending Platforms, Distributed Detection, and Adaptive Feedback," *Intel Technology Journal*, vol. 10, no. 4, 2006.

[11] D. Agrawal, K.-W. Lee, and J. Lobo, "Policy-Based Management of Networked Computing Systems," *IEEE Communications Magazine*, vol. 43, no. 10, pp. 69–75, 2005.

[12] J. Strassner, N. Agoulmine, and E. Lehtihet, "FOCALE: A Novel Autonomic Networking Architecture," in *Latin American Autonomic Computing Symposium (LAACS)*, 2006.

[13] N. Damianou, N. Dulay, E. Lupu, and M. Sloman, "The Ponder Policy Specification Language," in *International Workshop on Policies for Distributed Systems and Networks (POLICY)*, 2001.

[14] K. Twidle, N. Dulay, E. Lupu, and M. Sloman, "Ponder2: A Policy System for Autonomous Pervasive Environments," in *International Conference on Autonomic and Autonomous Systems (ICAS)*, 2009.

[15] L. Broto, D. Hagimont, P. Stolf, N. De Palma, and S. Temate, "Autonomic Management Policy Specification in Tune," in *Symposium on Applied Computing (SAC)*, 2008.

[16] NIST, "A Survey of Access Control Models," in *NIST Privilege (Access) Management Workshop*, 2009.

[17] S. De Capitani di Vimercati, S. Foresti, P. Samarati, and S. Jajodia, "Access Control Policies and Languages," *International Journal of Computational Science and Engineering*, vol. 3, no. 2, pp. 94–102, 2007.

[18] P. Loscocco and S. Smalley, "Integrating Flexible Support for Security Policies into the Linux Operating System," in *USENIX Annual Technical Conference*, 2001.

[19] M. Lacoste, T. Jarboui, and R. He, "A Component-Based Policy-Neutral Architecture for Kernel-Level Access Control," *Annals of Telecommunications*, vol. 64, no. 1-2, pp. 121–146, 2009.

[20] B. Lang, I. Foster, F. Siebenlist, R. Ananthakrishnan, and T. Freeman, "A Flexible Attribute-Based Access Control Method for Grid Computing," *Journal of Grid Computing*, vol. 7, no. 2, pp. 169–180, 2009.

[21] E. Damiani, S. Di Vimercati, and P. Samarati, "New Paradigms for Access Control in Open Environments," in *International Symposium on Signal Processing and Information*, 2005.

[22] OASIS, "eXtensible Access Control Markup Language (XACML)," 2010, http://www.oasis-open.org/.

[23] J. Ames, S. R., M. Gasser, and R. R. Schell, "Security Kernel Design and Implementation: An Introduction," *Computer*, vol. 16, no. 7, pp. 14–22, 1983.

[24] M. Aljnidi and J. Leneutre, "ASRBAC: A Security Administration Model for Mobile Autonomic Networks (MAutoNets)," in *4th International Workshop on Data Privacy Management (DPM) and Second International Workshop on Autonomous Spontaneous Security (SETOP)*, 2009.

[25] L. Badger, D. Sterne, D. Sherman, K. Walker, and S. Haghighat, "Practical Domain and Type Enforcement for UNIX," in *IEEE Symposium on Security and Privacy*, 1995.

[26] D. Bell and L. La Padula, "Secure Computer System: Unified Exposition and Multics Interpretation," MITRE Corporation, Bedford, MA, Tech. Rep. MTR-2997, 1975.

[27] D. Ferraiolo, R.Sandhu, S.Gavrila, D. Kuhn, and R.Chandramouli, "Proposed NIST Standard for Role-Based Access Control," *ACM Transactions on Information and System Security (TISSEC)*, vol. 4, no. 3, pp. 224–274, 2001.

[28] S. Marouf, M. Shehab, A. Squicciarini, and S. Sundareswaran, "Adaptive Reordering & Clustering Based Framework for Efficient XACML Policy Evaluation," *IEEE Transactions on Services Computing*, vol. 99, no. PrePrints, 2010.

[29] M. Serrano, S. van der Meer, J. Strassner, S. Paoli, A. Kerr, and C. Storni, "Trust and Reputation Policy-Based Mechanisms for Self-Protection in Autonomic Communications," in *International Conference on Autonomic and Trusted Computing (ATC)*, 2009.

[30] M. Anne, R. He, T. Jarboui, M. Lacoste, O. Lobry, G. Lorant, M. Louvel, J. Navas, V. Olive, J. Polakovic, M. Poulhiès, J. Pulou, S. Seyvoz, J. Tous, and T. Watteyne, "Think: View-Based Support of Non-Functional Properties in Embedded Systems," in *IEEE International Conference on Embedded Software and Systems (ICESS)*, 2009.

[31] A. Brown and C. Redlin, "Measuring the Effectiveness of Self-Healing Autonomic Systems," in *International Conference on Autonomic Computing (ICAC)*, 2005.

[32] C. Rippert, "Protection in Flexible Operating System Architectures," *Operating Systems Review*, vol. 37, no. 4, pp. 8–18, 2003.

[33] Cloud Security Alliance, "Top Threats To Cloud Computing," 2010, http://www.cloudsecurityalliance.org/topthreats.html.

[34] S. Berger, R. Cáceres, D. Pendarakis, R. Sailer, E. Valdez, R. Perez, W. Schildhauer, and D. Srinivasan, "TVDc: Managing Security in the Trusted Virtual Datacenter," *ACM SIGOPS Operating Systems Review*, vol. 42, no. 1, pp. 40–47, 2008.

[35] P. Ruth, J. Rhee, D. Xu, R. Kennell, and S. Goasguen, "Autonomic Live Adaptation of Virtual Computational Environments in a Multi-Domain Infrastructure," in *IEEE International Conference on Autonomic Computing (ICAC)*, 2006.

[36] J. Rutkowska and R. Wojtczuk, "The Qubes OS Architecture," Invisible Things Lab, Tech. Rep., 2010.

[37] T. Ristenpart, E. Tromer, H. Shacham, and S. Savage, "Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds," in *ACM Conference on Computer and Communications Security (CCS)*, 2009.

[38] R. He, M. Lacoste, J. Pulou, and J. Leneutre, "A DSL for Specifying Autonomic Security Management Strategies," in *Third IEEE International Workshop on Autonomous and Spontaneous Security (SETOP)*, 2010.

[39] J. Kephart and W. Walsh, "An Artificial Intelligence Perspective on Autonomic Computing Policies," in *IEEE International Workshop on Policies for Distributed Systems and Networks (POLICY)*, 2004.

[40] J. Strassner, J. de Souza, D. Raymer, S. Samudrala, S. Davy, and K. Barrett, "The Design of a New Policy Model to Support Ontology-Driven Reasoning for Autonomic Networking," in *Latin American Network Operations and Management Symposium (LANOMS)*, 2007.

Workshop-based Security Safeguard Selection with AURUM

Thomas Neubauer
Vienna University of Technology
Vienna, Austria
thomas.neubauer@tuwien.ac.at

Markus Pehn
SBA Research
Vienna, Austria
pehn@sba-research.org

Abstract - **Organizations are increasingly exposed to manifold threats concerning the security of their valuable business processes. Due to the increasing damage potential, decision makers are permanently forced to pay attention to security issues and are raising their security investments, but often (i) without considering the efficiency of the investments made, (ii) neglecting to involve people in order to raise security awareness and (iii) without full awareness of the importance of the decision at hand. This paper provides a crucial extension to the established risk management solution AURUM and extends its functionality by introducing the AURUM Workshop, which allows the selection of efficient safeguards based on corporate business processes. It highlights typical problems of (group) decision making and provides a solution to eliminate those shortcomings. Thereby, it supports decision makers in (i) refining the basic infrastructure elements to the specific requirements of the corporation, (ii) focusing on the most relevant risks and (iii) improving their awareness for the problem at hand.**

Keywords-Risk Management, AURUM, Decision Support

I. INTRODUCTION

Security hazards, such as viruses, hacker attacks or data theft, pose major threats to corporate assets and affect profit, shareholder value and a company's reputation. The increasing usage of the Internet leads to a rise in the frequency of security breaches related to information technology. Garg, Curtis and Halper [2] estimated security investments within US companies to reach about $30 billion by 2005. CERT estimated that about 90% of big and medium-sized companies were affected by security incidents in 2006. In May 2009, the New York Times reported on a billion dollar contract the US Government signed with security companies and universities with the aim of being equipped for so-called cyber warfare. Due to the continuous increase of information technology use and its monetary importance, the main questions posed by companies' managers are how to determine the optimum level of security investments and which measures are necessary and efficient.

This work provides an extension to the established risk management solution AURUM (AUtomated Risk and Utility Management; cf. [3], [4], [5], [6], [7], [8], [9]). AURUM provides a risk management solution that allows decision makers to evaluate security investments based on corporate business processes and infrastructure defined in a security ontology. A Bayesian network supports the risk definition, whereas an interactive multiobjective decision support approach is used for selecting safeguards. This paper extends the functionality of AURUM by introducing the AURUM Workshop. The AURUM Workshop provides the missing link between the ontology comprising corporate business processes and infrastructure, the Bayesian network, and the decision support module that allows the interactive selection of efficient safeguards. It takes typical psychological and social influence factors from literature into consideration. Thereby, it supports decision makers in (i) refining the basic infrastructure elements to the specific requirements of the corporation, (ii) focusing on the most important risks (risks with a high frequency, a high impact, or both) and (iii) improving their awareness for the problem (risks) at hand. The remainder of the paper is organized as follows. Section 2 introduces the state-of-the-art related to group decision making, whereas Section 3 gives a deeper insight into the psychological and sociological factors of group decision making. Section 4 introduces the AURUM Workshop. Sections 5 and 6 focus on describing the roles and the methods needed in the Workshop process. Finally, the Workshop process is described in detail in Section 7.

II. GROUP DECISION MAKING

Groups of persons are commonly employed in a various ways: to counterbalance individual subjectivity of goal and preference systems, assist creativity, compensate for complexity, and to increase members' identification with a decision (cf. [10], [11]). According to Frech [11], groups are similar to teams and characterized by "face to face contact of more than two persons over a longer time period oriented toward a goal identical to all members".

He argues that within a group, certain phenomena can be observed:

- A sense of togetherness also referred to as cohesion.
- The emerging of certain rules and restrictions in formal and informal interaction.

Staehle [12] uses these psychological facts to explain the difference between groups and teams. He defines a team as a focus-oriented work group with a stronger cohesion (team spirit) and stronger internal psychological relations. The role structure is more oriented to a team leader vs. member situation and the time span of working as a group is normally shorter. The designation of a team as "work group" leads to a point where teams are working toward a common goal, and individual preferences have to be shelved. This common goal leads to a higher level of cohesion in combination with high conflict potential. Groups bound by instructions,

characterized through the goal of achieving a common purpose, will be discussed below. Autonomic groups are not part of these considerations (cf. [13] for a definition and further discussion of autonomic groups).

A. Structure of Group Decisions

A group decision is the result of a group decision process (GDP). Laux [14] describes this process as two stages: the information process and the choice process (cf. [10], [13]). Paschka [15] divides the information stage into:

- Problem definition: Measure methods to elicit goal values. The problem definition can be predetermined by the management. A common approach is a comparison of actual and targeted business results that ends in a clear formulation of the problem's context.
- Detailed specification of the goal system: This means mainly the annulment of goal conflicts, approaches and intervention techniques in order to refine preference orders. A diversification of goal conflicts can be found in [16].

Paschka's choice stage consists of the following steps (cf. [14]):

- Determination of alternatives for action: These are multiple different methods, mostly selected via voting or exclusion approaches (or both), which also include decision finding based on these techniques (cf. [17]).
- Realization: Methods and techniques, for example project management to execute the decision.
- Control: Methods and techniques to compare the received results.

Once the information has been gathered, a discussion about possible alternatives and their results has to be carried out, usually leading to a voting process and a decision. This phase may possibly be influenced by members who try to further their individual preferences (cf. [14]). A participant of a group decision process is described through a set of variables (cf. [15]):

- Individual goal function and preference order: Depending on job position, knowledge, and interest in topic.
- Probability judgment: Depending on an individual's processing of given indicators, knowledge of the topic, and experience in similar situations.
- Information amount: Inside the group, external information will be presented through indicator values, but there is still the possibility that the amount of information certain group members have differs because of differences at the information processing level (prognostic function, cf. [14]) and external experience/knowledge (information structure) that is not available to the whole group (for example: secret strategic preferences of the management).

Based on these individual attributes it is obvious that, especially at the beginning of the information phase, each group member has different preferences and accordingly, a different preference order.

B. Application of the Group Decision Making Process to IS Safeguard Evaluation

Table I shows the application of the above described GDMP, according to Paschka and Laux, to information security by mapping the actions to the process.

TABLE I. APPLICATION OF IS SAFEGUARD EVALUATION TO THE GDMP

Phase of GDMP	Security Safeguard Evaluation Action(s)
Problem definition	Definition of cost and resource categories; Definition of tactical goals according to security policies.
Detailed specification of the goal system	Specification of strategic and tactical goals: analysis of the goal system and preference order (importance valuation of the goals) referring to the definition s made in the problem definition phase; Definition of assets, vulnerabilities, threats leading to risks.
Determination of alternatives	Definition of proper safeguards following the specified risks over a valuation scheme.
Realization	Implementation through physical, technical and administrative controls.
Control	IS control mechanism such as internal or external audits.

C. Structural Characteristics of Groups

Adler [18] describes cultural perspectives and background via a classification scheme:

- Homogeneous team/group: All members have the same cultural background
- Token team/group: All members expect one have the same cultural background
- Bicultural team/group: Two cultures that are represented by at least two members each
- Multicultural team/group: Three or more persons with different cultural backgrounds

Martirossian [19] describes homogeneous groups as more efficient for executing well-defined tasks, whereas more heterogeneous ones tend to find a greater number of feasible results. According to Adler [18], the monitoring effort increases with the degree of cultural difference. The problem solving approach as well as the communication mode can show large differences, which can be an opportunity but can also create risks in terms of misunderstandings and a lack of respect for personal attributes and behavior. The moderator can pick out the best of the available behaviors without harming group members, which can increase efficiency (cf. [18]).

Martirossian [19] argues that the *group size* is an important criterion. While big groups require a high degree of communication to include all members at a certain level, small groups are easier to handle in terms of communication, but bear risks like a lack of information or ideas. The workshop solution provided in this work tends to

involve a variety of different members, which requires good preparation and an experienced coordinator open to different problem solving structures.

Group leadership [19], which can be "people-oriented" and/or "goal-oriented", is an important criterion in a group. People-oriented leaders focus more on satisfying the group, while goal-oriented leaders place more emphasis on production and results. Both factors are important, as a balanced solution is recommended to hold a good Information Security Workshop. In a workshop situation, where the aim is to achieve optimal solutions, it is of utmost importance to structure the group with a view to the points described above. A certain degree of heterogeneity in team members' job positions (security experts as well as employees from outside the security field) and possibly their cultural background has to be handled with respect to balanced process leadership, which should be both goal and people-oriented to a certain degree.

III. PSYCHOLOGICAL AND SOCIOLOGICAL INFLUENCE FACTORS

Decision makers, no matter whether they act on their own or as part of a group, are usually confronted with a variety of psychological and social issues that have a major influence on their decisions (cf. e.g., [20]).

A. Basic Phenomena

1) *Confirmation trap*: Humans aspire towards consistency, which induces them to insist on the correctness of their actions and to ignore, eliminate or distort contrary information. Insist on belief effect: Works similarly to the confirmation trap mentioned above. Humans try to maintain their view of the world by ignoring, eliminating or distorting contrary information. Availability heuristic: Humans are able to remember some things better than others (cf. [21]). Possible reasons are emotional involvement, time, and spatial and sensory proximity [22], leading to an incorrect interpretation of these events by exaggerating their frequency, importance, etc. Anchoring and adjustment: The anchor is a basis for classifying new information based on a person's experience (cf. [23]). A lack of information often leads to the use of an arbitrary anchor, which causes a misclassification in relation to the anchor. Hindsight bias: After an event, people frequently believe that they predicted it correctly. There are a few theories concerning the origins of this mechanism:

- Relations were built after the event that do not or did not exist in reality.
- The theory of distorted answers (cf. [24]), which was formulated as a result of questioning eyewitnesses, shows that when people are confronted with irritant information, the capacity for remembering the facts decreases.
- The third theory is based on the abovementioned anchor heuristic, where the event is positioned too close to the anchor.

2) *Distortion by reasons of process variation*: People are generally inconsistent in their behavior. Lichtenstein and Slovic [23] as well as Tversky and Kahneman [25] have shown that this relation is not universally valid, and that logical procedure orientation and inductive behavior are only partially predetermined. Question structure: The formulation of the question is of vital importance to the processing and argumentation process inside respondents' minds.

3) *Prospect theory*: The frame in which a situation is embedded in terms of winning or losing dictates the expectations of the situation. If a loss is expected, a small benefit will be seen as a gain, whereas if a high benefit is expected, a small benefit will be handled as a loss (cf. [20]).

4) *Presentation of information*: Subjects are able to remember and categorize well presented information much better than badly presented information (cf. [20]).

B. Basic Phenomena in the Context of Group Decisions

The difficulty in mapping the basic phenomena to the group level is related to the nescience of specific group characteristics. In a group typically more resources, such as knowledge, power and financial capital are available.

- Availability heuristic at group level: Auer-Rizzi [20] takes it for granted that discussion of a prior case used as a prototypical example can affect the considerations in a positive or negative way.
- Anchoring at group level: Anchoring remains individual at group level; no group anchor is constructed, but rather individual anchors.
- Prospect theory at group level: Participants who see a situation as a gain are willing to take higher risks than others [25].
- Hindsight bias at group level: According to Stahlberg [26] there are no differences compared to the individual level if anchoring was used to provide the base of hindsight bias.

C. Influence of Majorities on Minorities and Vice Versa

The theory of social comparison (cf. [27]) postulates the human need to reassess own opinions. This mindset leads to behavioral uncertainty and the need for orientation towards reference points represented through

- a majority and the opinion it holds, or
- a strong individual opinion maker who persuades other participants of his view and, thereby, founds a majority.

According to Festinger [27] influencing majorities are one of the main reasons for distortion inside groups. An important factor within this theory is the divergence between physical and social reality. Physical reality is defined through the verifiability of facts, allowing everyone to check for themselves: e.g., financial data, statistics, etc. Social reality describes the common point of view

represented by a group or a strong majority. Asch [28] shows that in situations of divergence between social and physical reality, a tendency toward social reality is noticeable.

In contrast, Moscovici and Faucheux [29] showed that minorities can also influence the majority, if the minority argues with strong self-confidence and forcefulness. This refers explicitly to the behavior and not to fuzzy skill definitions (cf. [30]). In this case the majority tends to reflect on its point of view and often changes its opinion. Typical examples of this are influences on organizational hierarchies from outside the group structure. This arises for two reasons: A person who is accustomed to leading and can argue strongly also tends to be dominant within a group. Second, the behavior of subordinates is oriented towards their leader for reasons of personal benefit. This means expecting to gain favor by holding the view of the boss, and can occur consciously or unconsciously (cf. [31] for the theory of sociometric leader choice).

D. Readiness to Take Higher Risks at Group Level

People are ready to take higher risks at group level than in individual decisions (cf. examples [32], [33] and experiments [34], [35]).

- Allocation of responsibility: The risk level of group decisions increases with the number of liable participants (cf. [35]). A certain degree of anonymity arises as well, and risk aversion decreases with the degree of individual liability.
- A person who is willing to take higher risks has more influence: Individuals who tend towards risky decisions from the start argue more convincingly and are more successful at persuading others (cf. [36]).
- Social comparison: Brown [37] holds the view that risky decisions are preferred because of the social phenomenon that people willing to take higher risks have a better reputation. While this theory is not applicable in every situation, it often results in the unconscious attempt to take a little bit more risk than the other group members, which in turn leads to a positive evaluation of this person by the others.
- Strong arguments: According to Burnstein and Vinokur [38], group members are influenced by arguments that seem to be cogent, even in the case the argument or position being criticized is new and valid. Individual preferences as well as the characterization of the person and agreement with the person raising the argument can influence the rating of the argument.

E. Groupthink and its Criticism

Under certain circumstances, groups of sensible, smart, even shrewd men and women think and act in a way that can only be described as "collective stupidity" [32]. The most important psychological phenomenon in this area contains distortion mechanisms at individual and group level and results in a usually negative effect on decision finding.

The groupthink theory (cf. [39]) contains some preconditions that have to be met for groupthink to occur:

1) *High cohesion*: The phenomenon appears only in groups with a high or medium level of cohesion, due to the impossibility of individual members with a different view prevailing against the majority (cf. [40]). The main problem is the lack of disagreement and discussion in strong cohesive groups and the resulting isolation of opposition (cf. [20]).
2) *Compartmentalization* makes it easier to isolate oneself from external and new circumstances or restrictions. A compartmentalized group does not allow the influence of group harmony.
3) *Direct leading*: A patriarchic leader is not willing to accept disagreement.
4) The absence of *standardized decision procedures* leads to conformity and the loss of social control in the group's work.
5) Intragroup social and ideological homogeneity usually leads to homogeneous solutions of low impact due to the absence of opposition. The participants' goal is to achieve consensus at any price.
6) *Provocative and situational context*: Pressure on people with low self-esteem has a significant influence on the decision. Pressure to succeed leads to a high degree of conformity with group leaders' preferences. Low self-esteem arises from previous failures, excessive decision-making problems and moral dilemmas (cf. [40]).
7) *Tendency towards agreement*: People normally strive for harmony for reasons of conflict reduction within their environment.

Janis [33] characterizes the *symptoms of groupthink* with three, possibly overlapping, categories:

1) *Category 1 - overestimation of the group's own capabilities*: On the one hand, this is expressed by the illusion of invulnerability: the group holds the opinion that nobody and nothing is able to thwart them, which results in an extreme readiness to take risks. On the other hand, it results in the group's opinion that it upholds high moral and ethical standards, which creates a dilemma, as the group believes that everything it does is correct and, therefore, on a high ethical level.
2) *Category 2 – narrow-mindedness*: Everyone who holds a different view is excluded from the discussion. Further, stereotyping of opponents as well as collective resistance against warnings and different arguments is characteristic. Decisions that have already been taken are defended without considering new information and its implications.
3) *Category 3 - pressure toward uniformity*: This is mainly self-censorship that expresses itself in the tendency to keep doubts and misgivings to oneself.

Criticism of the groupthink theory is founded in part on the fuzzy definition of cohesion (cf. [40]). Classical conformity studies describe humans' aspirations towards a state of normative group conformity, where conformity within the group grows with an increasing degree of cohesion. Critics

address the case of different group norms: if the group norm does not prescribe the keeping of harmony but rather critical questioning, groupthink would be diminished. Furthermore, Janis holds the view that groupthink does not occur in groups with low cohesion. Schulz-Hart [40] disagrees and shows examples in which groupthink occurs in extreme situations or conditions such as compartmentalization outwards, homogeneity, directive leadership or extreme stress.

Janis only focuses on homogeneity of preferences, other forms are not accounted for. He does not explain how company-wide framework conditions can lead to unanimity. No methods are described for measuring low self-worth or hopelessness, and literature definitions disagree on contextual levels. In addition, there are some statements in Janis' work where the action-reaction relation is not sufficiently explained. Other issues include that there is no explanation of how overestimation of one's own capabilities and insularity can arise, because the pursuit of harmony or agreement cannot be used as an explanation [40]. There is also no clear specification what sort of consequences are caused by different preconditions. In response to the criticism Hart (cf. [32]) adds de-individualization. This concept is based on the work by Zimbardo [41] and defines unsocial, shortsighted behavior of groups and masses towards individuals with the goal of inducing groupthink. An opposing approach to explaining typical groupthink symptoms and the associated decision distortion was established by Whyte (cf. [42]), who bases his considerations on the prospect theory of Kahneman and Tversky [21]. Schulz-Hart [40] argues that the fiascos described by Janis [33] are founded on risk ignoring in case of loss expectancy and describes the groupthink characteristics as only cumulative values.

F. Decision Autism

According to Schulz-Hart [40] the distortion mechanism of decision autism occurs if a decision maker is controlled by self-affirmation tendencies. The symptoms are divided into 3 categories:

1) *Self-centered symptoms*: These are, first, a feeling of infallibility, which leads to a high degree of decision confidence and mental simplification of the problem area. Second there is self-reassurance, where any doubts are minimized by distorting them. The third is an increase in self-esteem, i.e., an increase of subjective confidence in oneself and one's opinion, combined with decreasing esteem for others and their opinions.
2) *Social symptoms*: Sniezek and Buckley [43] argue that social symptoms are not only of relevance within a group, but also for each individual, and lead to decision autism due to the fact that each individual acts in a social environment. Out of the whole range of social effects and symptoms, this refers to the ones who create selective communication. In this context, Schulz-Hart [40] has identified the following: more support for preferred discussion topics, selective attention, supporting likeminded people, and downplaying doubts. He also lists pressure on people who disagree, self-proclaimed mind guards and collective rationalization as symptoms resulting from personal attitudes that are only influenced at group level.
3) *Symptoms within the decision process*: Each step of the decision process potentially contains symptoms of decision autism (cf. [44]):

- Identification of the problem: Ignoring inconvenient problems, preference for supporting case studies.
- Generation of alternatives: Generation of fewer alternatives and focusing on the preferred one.
- Evaluation of alternatives: Distorted rating caused by selective information search, self-affirmation in evaluating information and rapid rejecting of divergent alternatives.
- Deciding: Lack of scrutiny of decisions.
- Implementation of the decision: Implementation without "what if" scenarios in mind.
- Control mechanisms: Excessive decision control.

IV. Overview of the AURUM Workshop

The AURUM Workshop is a process for supporting risk management. It is used to determine, refine and review security-relevant data needed as input for the AURUM risk management framework. The main characteristics of the AURUM Workshop are:

- Moderated: The workshop comprises three methods - Brainstorming/Discussion, Evaluation, and Selection - that are used by the moderator to get objective results from the workshop participants.
- Role based: Each process participant has a specific role that determines his tasks.
- Group based: Each process participant is a member of a small group of three or more people. By splitting one big group into several small groups, the approach aims to avoid psychological issues such as the "influence of majorities" and groupthink.
- Clear task structures: The process categorizes its tasks in three groups, where each is a basic type of task instances.
- Clear voting structures: The process provides a way to model consensus of opinions, which is based on the clear structures of the voting process.
- Awareness building: The AURUM Workshop aims to improve the security awareness of its participants in order to build an understanding of relevant risks, and options for their mitigation.

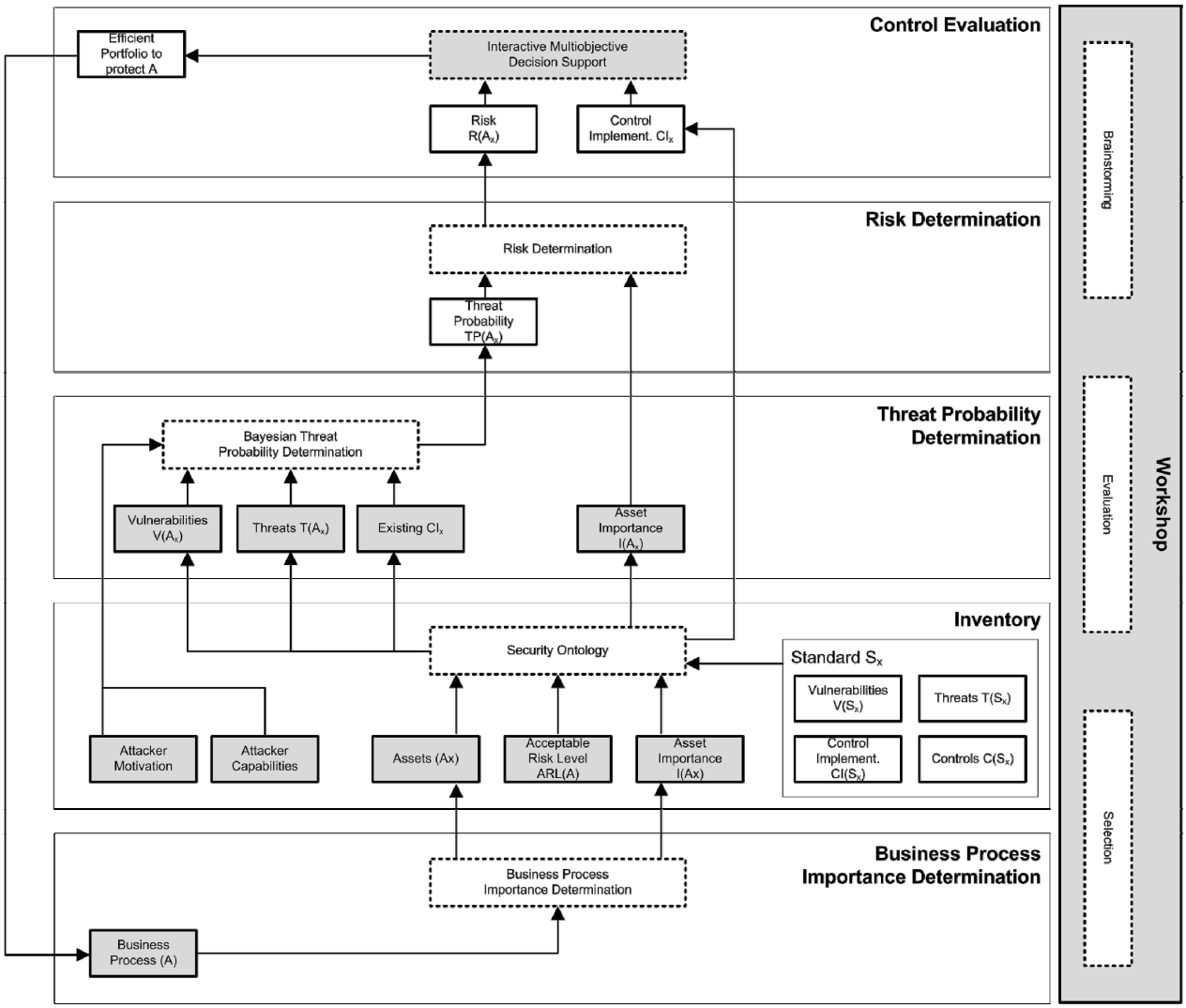


Figure 1. AURUM Workshop Process.

Figure 1 shows the overall scheme of the AURUM process and the integration of the AURUM Workshop. The gray squares denote activities and methods that are part of the workshop process. The workshop supports decision makers in going through the risk management process step by step. It supports the following risk management phases defined in AURUM: (i) Business Process Determination, (ii) Inventory, (iii) Threat Probability Determination and (iv) Control Evaluation. In the briefing phase, which is carried out prior to the workshop, the moderator selects a number of volunteers from different departments for the required roles.

V. ROLES

Heterogeneous group configuration can lead to impacts on decisions. To address this problem, this section outlines the roles used in the AURUM Workshop. A short description outlines each role's tasks and responsibilities, followed by a list of recommended skills for each role. We describe the main tasks of the role during the workshop and the interaction with other process participants (specifically whether and how they exchange knowledge and experience for task execution).

A. Moderator

Short Description: Ideally, a security consultant familiar with the AURUM process and the business area should be selected as moderator. It is highly recommended that the role of the moderator is assumed by an external consultant familiar with the process and its typical problems. If there is a high number of participants, two moderators can be selected.

Recommended Skills: High familiarity with AURUM, the ability to nurture security awareness and build an understanding of security in the participants' minds,

consideration of psychological issues that can occur in group decision processes.

Main Tasks and Interaction: The moderator has one of the main roles in the AURUM Workshop. He has to administrate groups and data input: (i) He defines small groups, creates user accounts and assigns roles to all process participants. (ii) Data exchange Data input tasks are mostly brainstorming, which involves naming and the need for merging and deleting. (iii) Data input: The moderator can choose whether he wants to join one of the small groups. In that case, he must input data like all other process participants. Due to the complexity and number of tasks the moderator has to deal with, this is not recommended. (iv) Regulation of interaction: The AURUM Workshop is characterized by highly interactive tasks where the moderator is the main interaction controller: he opens the tasks, users input data and discuss, and then he closes the task.

Involvement in Group: The moderator leads and guides the participants through the workshop. In this respect it is also not impossible for the moderator to be a member of a group, but because of the number and complexity of the tasks it is not recommended. Additionally, prior knowledge and his position of respect with regard to the other process participants can cause group dynamic issues.

B. Management Member

Short Description: This role is ideally assumed by members of middle or high level management who contribute to the group with structural and process knowledge. It must be taken into consideration that the dynamics between majorities and minorities within the group can cause problems with established authority relations outside the group. Therefore, it should be attempted to form groups at the same or similar levels of hierarchy. The presence of management members is an indispensable cornerstone of an accepted process execution at management level, due to their knowledge of cost restrictions and running processes. Before process execution, it is essential to ensure management support and therefore sufficient presence of management members.

Recommended Skills: (i) Process knowledge: Management members have to be aware of the strategic goals of internal or external business processes so as not to lose sight of integration problems that could be caused by new security controls. (ii) Knowledge of cost structures: The best security control set is worthless if it cannot be implemented due to cost restrictions. Including cost considerations from the beginning can eliminate unrealistic economical security control estimations. This requires a high degree of interaction with members in "security" roles concerning possible safeguard costs, and therefore security members with some knowledge of costs. (iii) The ability to present decisions and their costs at management level is indispensable for the adoption of the developed solutions. To build this understanding in management members' minds, it is necessary to impart knowledge about effects of security incidents before process execution. (iv) The task "category voting", in particular, has to be executed under guidance of the group's management member, who should have knowledge about cost and value categories the company uses and the ability to give the other process participant an idea of these categories..

Main Tasks: A rating task is done by the individual user and indicates an assignment of numerical values that were voted upon. Of special interest for management members is the task category of voting, where they contribute with data input as well as performing the leading role because of their special knowledge about cost and process structures.

Involvement in Group: Each management member is directly integrated in exactly one group. The interaction with members of other groups happens through discussion tasks. Each management member is involved in all group decisions and has the leading role in category evaluation.

C. Expert Member

Short Description: Depending on the problem area, an expert member can come from a different department. In our considerations these mostly concern information security employees, but generally this could be any department that tries to evaluate security claims and corresponding safeguards. An expert member fills the gap between the structural and cost knowledge of management members and the user experience of the key process user. Other process participants improve the expert's knowledge by giving him a broader view of the issues. Security expert members are essential for identifying the problem space and ideally help to understand problems at other business levels. It is highly recommended that the importance of team-oriented work is borne in mind in selecting the expert member. Expert members who are not willing or able to share knowledge and responsibility are a destructive power and impede the process.

Recommended Skills: (i) Infrastructural knowledge to handle the asset identification is one of the main skills an expert member must have. This also holds true for experience with past incidents affecting the assets in question and their occurrence rates, which is essential for refining probability ratios. (ii) Technology knowledge: The ability to identify and estimate possible synergy effects and effectiveness of safeguard candidates. Records and statistics can be of additional help in these steps. (iii) Cost knowledge, which is important in interacting with the management members. Without feasible estimations about possible safeguard implementation costs, the management is unable to consider cost restrictions in the evaluation process. It is rarely possible to give even a rough estimate because of the difficulty of determining issues like installation, maintenance, etc.

Main Tasks: A rating task is performed by an individual user and is an assignment of numerical values to tasks that were voted upon. Of special interest for expert members is the task asset voting, where they input definition data because of their special infrastructural knowledge, as mentioned above. Risk voting is also performed only by the expert members and deals with the estimation of occurrence rates.

Involvement in Group: Each expert member is directly integrated in exactly one group. The interaction with members of other groups happens through discussion tasks as described above. Each expert member is involved in all group decisions.

D. Key Process User

Short Description: In addition to the structural and expert knowledge provided by management and expert members, users must also contribute to the process. The participation of key process users also enhances acceptance of the decided actions and their costs at employee level. In the selection of key process users, it is recommended that their prior knowledge and openness to new approaches are taken into consideration. Candidates with negative attitudes towards new ideas can cause major acceptance problems. It is essential to understand that if the key process users are not convinced of the approach, they will not be able to communicate the idea and need for information security measures at their business level.

Recommended Skills: (i) Experience with main business processes and tasks to enable use of user know-how within the process considerations. This, in particular, takes data input problems into consideration at the task execution level, as well as experience with the use of previous information security measures. (ii) Ability to defend unwelcome decisions at execution level founded on in-depth knowledge about their necessity. This is based upon the introduction to security problems and possible consequences at the briefing held prior to the process.

Main Tasks: A rating task is done by the individual user and indicates an assignment of numerical values to voted tasks. In the current version this is only done manually to get an asset ranking and to evaluate the incident occurrence rate; the rating of the other steps occurs automatically via the number of mentions. Discussions are lead by the moderator, who is the only one who can perform data changes on that basis. Unlike the other roles, the key process users themselves do not have any tasks in which they assume a leading role.

Involvement in Group: Each key process user is directly integrated in exactly one group. The interaction with members of other groups happens through discussion tasks as described above. Each key process user is involved in all group decisions.

VI. WORKSHOP METHODS

The workshop comprises three methods that are used by the moderator to generate data necessary to carry out the risk management process (see Table II): Brainstorming, Evaluation, and Selection.

1) *Brainstorming:* Brainstorming enables a group of decision makers to quickly assess the data relevant for the information security of their organization. The system supports the decision makers as they enter as many items as they judge appropriate.
2) *Evaluation:* Based on Grünbacher (cf. [45]) we use a border criterion voting mechanism for rating the items gathered during brainstorming. Each participant decides upon the importance and ease of implementation of the so-called win conditions. The system calculates a medium value depending on the degree of consensus.
The voting results are underlined with a traffic light system to signal the degree of controversiality using the colors red (<50% consensus), orange (>=50 and <=75% consensus) and green (>75% consensus). The borders are variable and arise from task-dependent mathematical methods: (i) Taking numerical values as input, the standard deviation of the input values from the different decision makers is used to determine the threshold and, thus, the degree of consensus. (ii) Taking the number of votes as input, the number of votes related to the total number of voters determines the threshold and, thus, the degree of consensus. To avoid disagreement, e.g., out of ignorance, the voters are instructed not to vote if they do not know. The evaluation process can be summarized as follows:
   a) A set of possible win conditions arising from brainstorming phases are the input for voting.
   b) Each possible win condition is voted on in the categories of business importance and ease of realization. To avoid distortion from blind votes, the members are instructed not to vote if they do not know.
   c) The average of each condition over the two categories will be displayed, and a red/green colored marking indicates the degree of consensus.
   d) A structural discussion helps to clarify reasons for disagreement and convey tactical knowledge known to individuals to the rest of the group [45].
3) *Selection/Discussion:* During a group discussion based on the ratings' analysis, the group decides which items are to be selected. If judged necessary, the brainstorming and rating steps can be repeated. Discussion tasks have to be carried out after voting in order to resolve any disagreements. The degree of consensus or disagreement an issue receives determines how it is handled in the discussion, with the moderator acting as a mediator. Of course, the nature of group discussions is always to some extent undefined and it is difficult to determine concrete rules. Therefore, a moderator with high psychological and didactic competence is required. Nevertheless, some general suggestions are made below to aid the moderator in this complicated task.

It is suggested to address orange and red color items by questioning:

- Ask an individual member why he or she thinks that a point is important or not. The points that he or she mentions will certainly be agreed or disagreed with by several members, forming the basis for the discussion.
- Allow constructive interruptions but make sure to guard against domination by a few members (cf. Chapter 4.5, especially the problem of "majorities and minorities").
- If only one person has mentioned a specific issue, do not ask this person why he or she thinks that it is important. Instead ask another member in order to avoid the human

tendency to wait for explanations (cf. Chapter 4.5); possibly he or she will bring up issues nobody has thought about.

In the end a generally accepted solution/rating should be found. If this is not possible after an adequate amount of time, the only possibility for the moderator is to overrule the disagreeing parties with a compromise. This must be considered a last option and should be avoided whenever possible.

TABLE II. CHARACTERISTICS OF TASK TYPES

	Brainstorming Task	**Evaluation Task**	**Selection/Discussion Task**
Executors	Participants in a specific role (instance dependent).	Participants in a specific role (instance dependent).	All participants.
Input	List of issues which have to be rated mathematically.	The question what is imaginable for ... ; i.e., a brainstorming request.	A list of issues from a prior voting task, containing items on which the participants do not agree or agree only in part.
Output	List of numerical values assigned to issues.	A list of written issues with a certain degree of consensus.	Changes in the input list which represent a more accepted output list, and/or more sophisticated user.

VII. THE AURUM WORKSHOP PROCESS

This section explains the phases of the AURUM Workshop in detail. Each step is described according to the three criteria of input, output, and sub-steps. The sub-steps list the necessary internal tasks and explain the reason and the type of task for each one, breaking down a quite complex process step into manageable and understandable topics.

A. Workshop Briefing

After reviewing and selecting workshop participants according to their profile and the requirement definition in section 7.2, the members are briefed on the goals of the process: (i) Definition of the risk analysis context and goals: This first step aims at defining the scope of the workshop and its goals. This is required for the orientation of the process and the definition of criteria and to measure its success. (ii) Selection of workshop participants: In order to raise the efficiency of the workshop session in terms of quality and quantity of the workshop output, the moderator must select participants according to their knowledge, their suitability. Participants should be selected to cover the whole spectrum of security problems and include a manager in charge of the decisions to emerge from this process. (iii) Psychological issues: With knowledge of psychological dynamics in group decisions, the participants may be able to avoid typical problems. (iv) AURUM Workshop process: Participants are informed about the process steps, especially input and expected output data. This has to happen in a way that ensures the members understand their roles and, therefore, their integration in the process, including issues such as voting mechanisms, group structuring, etc. (v) Terminology: For successfully conducting the workshop part of the process, it is essential to impart knowledge of basic security terms and how they relate.

The following section outlines suggestions for briefing the workshop participants, especially concerning issues arising from related work. The main points of concern during the execution of the process are the following:

- Why the workshop is carried out: Which goals and prospects regarding process output exist and how this approach differs from previous ones.
- Characterization of the business unit: The participants have to be informed about affected business. If most participants (especially non expert members) have only limited knowledge about the field in question, a short introduction is suggested. It is assumed that sufficient knowledge about business concerns will help to understand problem complexity and needs. In case of a general information security safeguard evaluation this point can be ignored, focusing more generally on process goals.
- Explaining MOSEP workflow: Participants have to be informed about the individual steps (cf. chapter 8.6); in particular, input and expected output data are important for seeing the overall picture in terms of risk assessment (cf. chapter 3).
- Understanding of security terms and their mashing: It is essential for performing the workshop part of the process to impart knowledge of basic security terms (cf. chapter 2) and how they relate.
- Building security awareness: The awareness problem was discussed in chapter 6. Participants have to understand difficult terms, like "social engineering", "human asset" etc., to perform a more feasible evaluation.
- Building awareness of possible psychological influence: With knowledge of psychological dynamics in group decisions (cf. chapter 4.5), the participants may be able to avoid typical problems.

The participants are asked to perform an interactive knowledge exchange through question/answer interaction. Each moderator uses different methods of interaction and communication, depending on personal experience and preferences.

B. Phase 1: Business Process Importance Determination

Description: This step aims to identify the most relevant business processes. For this purpose, the expert group is asked to execute a brainstorming and evaluation task. Gross discrepancies (foremost red-colored items) have to be discussed by the workshop members, and result in an accepted list of processes ranked by their importance.
Steps:

- Business Process Selection: The decision makers select the business processes to be evaluated. This step includes the discussion of the selected processes and their ranking in the event of a low degree of consensus. In order to resolve such problems, the moderator should discuss the following questions with the workshop participants: "Why were certain processes mentioned?" and "Why did certain members vote high and others low for the importance of an issue?"
- Business Process Importance Determination: The decision makers determine the importance of the selected business processes within the corporation, and their need for protection.

Main Question: What should be protected?
Output: An accepted list of business processes ranked by importance.

C. Phase 2: Inventory

Description: This step aims to identify the most relevant assets. For this purpose, the expert group is asked to execute a brainstorming and evaluation task. Gross discrepancies have to be discussed by the workshop members, and result in an accepted list of assets ranked by their importance. Note that this phase can be supported by the AURUM security ontology, which already contains a wide selection of assets. Thus, decision makers only need to review the assets proposed by the ontology and the discussion can focus on the issues where little consensus exists.
Steps:

- Assets: This step includes the discussion of the assets corresponding to the selected processes.
- Asset Importance Determination: The decision makers determine the importance of the selected assets, and, thus, their need for protection. The decision makers can use a suggestion made by the system that is calculated based on the importance of the business processes (cf. [6]).
- Acceptable Risk Level: Level of risk judged to be outweighed by corresponding benefits or one that is of such a degree that it is considered to pose minimal potential for adverse effects.
- Attacker Capabilities: This step aims to evaluate and define the capabilities of potential attackers.
- Attacker Motivation: This step aims to evaluate and define the motivation of potential attackers.

Main Question: Which assets exist, and which of them are really worth protecting?

Output: An accepted list of assets ranked by their importance, the acceptable risk level for each business process, the attacker capabilities, and the attacker motivation.

D. Phase 3: Threat Probability Determination

Description: This step aims to determine and review vulnerabilities, threats and existing countermeasures. It aims to evaluate possible threats and their causes. The basic data for this purpose is the asset list assembled in process step 1. First, the possible threats for each asset have to be determined, which happens through group voting. The result is a list of threats, in which each threat has to be argued by listing dangers (also group voting), which produces a list of vulnerabilities for each threat. The vulnerability and the threat determination have to be concluded by a discussion task based on the degree of consensus in the two voting steps. For this purpose, the expert group is asked to execute a brainstorming and evaluation task. Gross discrepancies have to be discussed by the workshop members, and result in an accepted list of vulnerabilities and threats ranked by their importance. Note that this phase can be supported by the AURUM security ontology, which already contains a wide selection of vulnerabilities and threats based on established security standards such as ISO 27001 or NIST SP 800. Thus, decision makers only need to review the assets proposed by the ontology, and discussion can focus on the issues where little consensus exists. In this case voting can be limited to selection tasks, the vulnerabilities follow automatically and only have to be adapted to the specific business needs.
Steps:

- Vulnerabilities: Based on the list of threats, the next step deals with determining the causes for each threat.
- Threats: This sub-step attempts to evaluate a set of corresponding threats for each asset. The execution as voting task requires brainstorming on behalf of the group and input concerning problematic circumstances. The moderator aggregates the data to obtain the list of threats for each asset that is the output of this sub-step.
- Existing countermeasures: This step aims to review and evaluate existing countermeasures.

Main Question: Which dangers are the individual assets exposed to?

Output: Accepted lists of threats and corresponding vulnerabilities.

E. Phase 4: Control Evaluation

Description: Based on the risk evaluation, the set of possible administrative, technical and physical controls required to avoid such incidents must be determined. This is achieved by voting, followed by a discussion. The output is a set of controls for each risk. Alternatively, it is possible to define only the requirements for control. Concrete products can be determined in the post-workshop evaluation step.
Steps:

- Criteria Definition: This step defines a set of criteria concerning business conditions and possibly related enterprise-wide controlling mechanisms.
- Interactive Selection: This step supports decision makers in determining the solution that best fits their ideas and objectives, choosing from the possibly hundreds (or even thousands) of Pareto-efficient alternatives of countermeasure portfolios identified previously. The procedure starts with an efficient portfolio and allows the decision maker to iteratively move in solution space towards more attractive alternatives until no "better" portfolio can be found. The system provides immediate feedback about the consequences of different choices in terms of the remaining alternatives and, thereby, allows the decision maker to evaluate different investment scenarios. The system provides the decision maker with ample information on the specific selection problem and ensures that the finally selected solution will be an optimal (i.e., Pareto-efficient) one.

Main Question: Which countermeasures are possible?
Output: Accepted lists of countermeasure portfolios for protecting the selected business processes.

VIII. CONCLUSION

Managers regularly have to cope with a wide spectrum of potential risks and, therefore, the decision of selecting the most appropriate set of security safeguards. Moreover, they are challenged by legal and economic requirements leading to the demand to carry out risk assessment on a regular basis. This paper proposed an approach called AURUM Workshop for integrating the advantages of workshops into the established risk management solution AURUM. It provides decision makers with a stepwise method for risk assessment by taking into account and mitigating typical psychological and social influence factors that usually occur in (group) decision processes. Decision makers are supported by a moderator who provides professional advice during the entire process and reduces the influence of individual opinions on the whole decision. AURUM Workshop is intended to not only evaluate data, but also to impart security awareness to the participants in order to build an understanding of relevant risks, and options for their mitigation. It supports decision makers in identifying and focusing on the most important risks and provides intuitive interactive decision support for evaluating different protection scenarios.

ACKNOWLEDGMENT

This work was performed at the Vienna University of Technology and the research center Secure Business Austria funded by the Federal Ministry of Economy, Family and Youth of the Republic of Austria, and the City of Vienna.

REFERENCES

[1] Workshop-Based Risk Assessment for the Definition of Secure Business Processes; Thomas Neubauer and Markus Pehn; International Conference on Information, Process, and Knowledge Management (eKNOW'10), IEEE Computer Society, 2010, pp. 74-79.

[2] A. Garg, J. Curtis, and H. Halper, "Quantifying the financial impact of it security breaches," *Information Management & Computer Security*, vol. 11/2, 2003, pp. 74–83.

[3] A. Ekelhart, T. Neubauer, and S. Fenz, "Automated risk and utility management," in *2009 Sixth International Conference on Information Technology: New Generations*. IEEE Computer Society, 2009, pp. 393–398.

[4] A. Ekelhart, S. Fenz, and T. Neubauer, "Ontology-based decision support for information security risk management," in *International Conference on Systems, 2009. ICONS 2009.* IEEE Computer Society, March 2009, pp. 80–85.

[5] ------, "Aurum: A framework for supporting information security risk management," in *Proceedings of the 42nd Hawaii International Conference on System Sciences, HICSS2009*Los Alamitos, CA, USA: IEEE Computer Society, January 2009, pp. 1–10, 978-0-7695-3450-3.

[6] S. Fenz, A. Ekelhart, and T. Neubauer, "Business process-based resource importance determination," in *Proceedings of the 7th International Conference on Business Process Management (BPM'2009)*. Springer, 2009, pp. 113–127.

[7] T. Neubauer and C. Stummer, "Extending Business Process Management to Determine Efficient IT Investments," in *Proceedings of the 2007 ACM Symposium on Applied Computing*, 2007, pp. 1250-1256.

[8] T. Neubauer, A. Ekelhart, and S. Fenz, "Interactive selection of ISO 27001 controls under multiple objectives," in *Proceedings of the Ifip Tc 11 23rd International Information Security Conference, IFIPSec 2008*, vol. 278/2008. Boston: Springer, July 2008, pp. 477–492.

[9] T. Neubauer and C. Stummer, "Interactive selection of web services under multiple objectives," *Information Technology and Management,* vol. 11(1), 2010, pp. 25-41.

[10] E. Kahle, *Betriebliche Entscheidungen*Oldenburg, vol. 6, 2001.

[11] M. Frech, Arbeit in und mit Gruppen in Kasper, H. and Maierhofer, W.(eds.) Personalmanagement-Führung - Organisation. Wirtschaftsverlag Ueberreuter, 1996.

[12] W. Staehle, Management - Eine verhaltenswissenschaftliche Perspektive. München, 1991.

[13] E. Saliger, Betriebswirtschaftliche Entscheidungstheorie. Oldenburg, 2003.

[14] H. Laux, *Entscheidungstheorie*. vol. 6 Springer, 2007.

[15] R. Paschka, *Multipersonalität bei Mehrfachentscheidungen*. Deutscher Universitätsverlag, 1995.

[16] J. Bidlingmaiyer, *Unternehmerische Zielkonflikte und Ansätze zu ihrer Lösung*. Zeitschrift für Betriebswirtschaft, 38, No.3, 1968, pp.149 - 179.

[17] B. Roy, Decision Aid and Decision Making in Costa, C. A. Bana e (ed.): Readings in Multiple Criteria Decision Making. Berlin, 1990.

[18] K. Adler, International Dimensions of Organizational Behavior 4th Edition. Ohio, South Western/Thomson Learning, 2002.

[19] J. Martirossian, *Decision Making in Communities: Why Groups of Smart People Sometimes Make Bad Decisions*. Community Association Press, A Division of Community Association Institute, 2001.

[20] W. Auer-Rizzi, Entscheidungsprozesse in Gruppen - kognitive und soziale Verzerrungstendenzen. Wiesbaden, DUV, 1999.

[21] A. Tversky and D. Kahneman, "Availability: A heuristic for judging frequency and probability." *Cognitive Psychology*, vol. 5, 1973, pp. 207–232.

[22] R. Nisbett and L. Ross, *Human Inference: Strategies and Shortcomings of Social Judgment.* Englewood Cliffs: Prentice Hall, 1980.

[23] P. Slovic and S. Lichtenstein, "Comparison of Bayesian and Regression Approaches in the Study of Information Processing in Judgment," *Organizational Behavior and Human Performance*, vol. 6, 1971, pp. 649 – 744.

[24] M. McCloskey and M. Zaragoza, "Misleading postevent info and memory of events: Arguments and evidence against memory impairment hypothesis." *Journal of Experimental Psychology: General*, vol. 114, 1985, pp. 1 – 16.

[25] A. Tversky and D. Kahneman, "The framing of decisions and the psychology of choice," *Science*, vol. 211, 1981, pp. 453 – 458.

[26] D. Stahlberg, A. Maas, and D. Frey, "We knew it all along: Hindsight bias in groups." *Organizational Behavior and Human Decision Processes*, vol. 63, 1995, pp. 46–58.

[27] L. Festinger, "Informal social communication," *Psychol. Rev.*, vol. 57, 1950, pp. 271–282.

[28] S. Asch, "Studies of independence and conformity: a minority of one against an unanimous majority." *Psychol. Monogr.*, vol. 70, 1956, p. 9.

[29] S. Moscovici and C. Faucheux, *Social influence, conformity bias and the study of active minorities*, t. E. In: Berkowitz, L. Advances in experimental social psychology, Ed. Academic Press, New York-London, 1972.

[30] R. D. Mann, "A review of the relationships between personality and performance in small groups," *Psychol. Bull.*, vol. 56, 1959, pp. 241–270.

[31] R. F. Bales and P. Salter, "Role differentiation in small decision making groups" in Parsons, T. and Bales, R. F.: "Family, socialization, and interaction process," *Free Press, Glencoe/Illinois*, vol. 1, 1995, p. 1.

[32] P. Hart, *Groupthink in government: A study of small groups and policy failure*. Amsterdam, Swets & Zeitlinger, 1990.

[33] I. Janis, *Groupthink. Psychological Studies of Policy Decisions and Fiascoes*. Boston: Houghton Mufflin, 1982.

[34] J. Stoner, "A comparison of individual and group decisions involving risk." Ph.D. dissertation, School of Industrial Management, M.I.T., 1961.

[35] M. Wallach, N. Kogan, and D. Bem, "Group influence on individual risk taking." *J.Abnorm. Soc. Psychol.*, vol. 65, 1962, pp. 75–86.

[36] B. Collins and H. Guetzkow, A social psychology of group processes for decision making. Wiley, New York, 1964.

[37] Brown, R. (1965). Social psychology. New York: Free Press.

[38] E. Burnstein and A. Vinokur, "Testing two classes of theories about group-induced shifts in individual choice." *J. Exp. Social Psychology*, vol. 13, 1973, pp. 315–332.

[39] H. Franke, Problemlösen in Gruppen: Veränderungen im Unternehmen zielwirksam realisieren, 3, Ed. Leonberg:Rosenberger Fachverlag, 1998.

[40] S. Schulz-Hart, Realitätsflucht in Entscheidungsprozessen - von Groupthink zum Entscheidungsautismus. Ber: Verlag Hans Huber, 1997.

[41] P. Zimbardo, "The human choice: Individuation, reason and order versus deindividuation, impulse and chaos," *In: Arnold W. J. and Levin D. (Eds.): Nebraska symposium on motivation, University of Nebraska Press, Lincoln*, vol. 17, 1969, p. 1.

[42] Whyte, G., & Sebenius, J. K. (1997). The effect of multiple anchors on anchoring in individual and group judgment. Organizational Behavior and Human Decision Processes, 69, 75–85.

[43] J. Sniezek and T. Buckley, "Cueing and cognitive conflict in judge-advisor decision making." *Organizational Behavior and Human Decision Processes*, vol. 62, pp. 1995, 159–174.

[44] R. Aldag and S. R. Fuller., "Beyond fiasco: A reappraisal of the groupthink phenomenon and a new model of group decision processes," *Psychological Bulletin*, vol. 113, 1993, pp. 533–552.

[45] P. Gruenbacher and R. Briggs, "Surfacing tacit knowledge in requirements negotiation: Experiences using EasyWinWin," *Proceedings of the 34th Hawaii International Conference on System Sciences*, vol. 34, 2001, pp. 1–8.

An Holistic Approach to Public/Private–Key Based Security in Locator/Identifier–Split Architectures

Oliver Hanka
Technische Universität München
Institute of Communication Networks
80333 Munich, Germany
oliver.hanka@tum.de

Wolfgang Fritz
Leibniz Supercomputing Centre
85748 Garching, Germany
wolfgang.fritz@lrz.de

***Abstract*—Network security has become an essential business requirement over the past few years. As this demand will increase even more in the future, researchers agree that security must be a key element for any novel Next Generation Internet architecture. Contrary to today's add–on approach to security, the mechanisms must be anchored in the overall architecture and should be a major concern already during the design phase. In this article we present an approach based on the private/public–key principle for almost any locator/identifier–split architecture. We suggest to extend the mapping system to also serve as public–key infrastructure and recommend to use smart cards for the client side key management.**

***Keywords*-Public–Key Infrastructure, Locator/Identifier–Split, Smart Cards, Asymmetric Cryptography, Next Generation Internet, HiiMap**

I. Introduction

Today's Internet architecture faces some well-known limitations and many ongoing research activities exist to define the so called *Next Generation Internet* (NGI). For example, there is currently only one address representing different aspects—the IP address stands both for the particular host we want to contact and for the topological location, where it can be reached. That's why many clean–slate approaches towards an NGI architecture favor a so called locator/identifier–split [1][2][3][4][5][6]. Furthermore, as the Internet evolved during the years, many new aspects had to be considered, i.e., how to communicate in a secure way? Many small and different solutions have been applied to the architecture to answer this question. Rather than using these numerous add-ons, it is agreed that security needs to be an integral part in future concepts, providing an holistic approach to guarantee secure communication. This is because the Internet has transformed from a communication means to transfer files and messages between some few nodes to the basis of today's economy with billions of participants.

Like Moskowitz et al. [4], for example, many have suggested linking the identifier with a public–key in some way. This has the benefit that each communication partner can be authenticated based on the public/private–key principle by Diffie et al. [7]. Additionally, it can be used to exchange a symmetric secret for stream encryption.

Moskowitz et al. suggest hashing the public–key and using it as the identifier of that node. This, however, raises some problems as described in [2]. For example, it it easy to find a random private–key, public–key, identifier triple and furthermore, the public–key can not be exchanged while keeping the identifier. Therefore, we propose a *loose coupling* between the public–key and the identifier [2]. For that loose coupling, the relation between a certain key and an identifier is stored in the mapping system.

In this article we introduce a way to use a mapping system of an locator/identifier–split architecture as a public–key infrastructure. One very important aspect of the private/public–key principle is the integrity of the public–key. Therefore, we focus on the retrieval of the public–key from the mapping system and discuss how the user can verify the integrity of it. Additionally, we describe the key management on the client side supported by smart cards. We detail the initial bootstrap process and discuss the mechanisms for an encrypted communication. We will also consider client devices with low computational power, like sensors, as they already play an important role today that will even more increase in the future.

The remainder of this article is structured as follows: Section II discusses related work like the Host Identity Protocol, some key features of UMTS/GSM regarding security and basics of a public–key infrastructure. In Section III, we give a brief overview of HiiMap which we use as an example architecture throughout the rest of this article. Afterwards in Section IV we detail our approach to integrate the public–key infrastructure into the mapping system. The smart card based client key management is outline in Section V. Before we conclude our work in Section VII, we will evaluate the concept in Section VI.

II. Related Work

In the following, we will first give a brief introduction of the so called locator/identifier–split principle and discuss one example architecture *(Host Identity Protocol (HIP))* that

is based on it. Afterwards, some key functionalities of the *Universal Mobile Telecommunications System*'s (UMTS') security mechanisms are described, as some approaches are similar to the ones used in our concept (see Section V). Finally, we outline the key concepts of today's *Public Key Infrastructure* (PKI).

A. Locator/identifier-split

In today's Internet, the IP–address represents in fact two different meanings: First, it is of course an identifier of the particular node we want to contact (*who?*). Secondly, it also answers the question how this node can be reached (*where?*). Many *Next Generation Internet* (NGI) approaches propose to use a so called locator/identifier–split, in order to answer only one question at a time. In these concepts, there is the *identifier*, that stands for the endpoint we want to contact. In contrast, the *locator* answers the question how this can be done. Therefore, we have two different addresses, one for each meaning. By providing these, the locator/identifier–split solves several issues concerning mobility, routing table growth and scalability. Additionally, it leaves some space to integrate security mechanisms into the network layer.

B. Host Identity Protocol

HIP [4] uses the *Host Identity Tag* (HIT) as identifier. The HIT either represents the 128 bit long public key or—in case of greater key length—the hash of it. In that way, any node can verify the public key of its peer by only knowing the identifier. Therefore, no PKI is required. During HIP base exchange, the public keys and a secret are exchanged [8].

As already stated in the introduction, this approach has a major security vulnerability. An attacker could start to generate many private/public–key pairs and hash the public key into a HIT. In a next step, he could query the mapping system and check whether the HIT is already reserved. In case he finds an already reserved HIT, the attacker holds a valid private key to that HIT. This does not enable an attacker to find a specific private key for a certain HIT, but allows for random attacks and could become interesting for Botnets, for example.

C. UMTS/GSM

The *Global System for Mobile Communications* (GSM) has some major drawbacks as described in [9][10]. As it was not initially designed for Internet purposes, it faces many new challenges. Most of them are eliminated in pure UMTS environments [11][12]. Nevertheless, some problems still exist when roaming from UMTS to GSM and vice versa is supported [13].

The UMTS architecture also uses a smart card based principle for authenticating its clients—the so called *Authentication and Key Agreement* (AKA) [13]. In addition to authentication, it provides data encryption (with cipher key CK) as well as integrity protection (with integrity key IK). Similar to parts of our concept, UMTS is therefore able to prove the integrity of received messages. In contrary, UMTS security functions only protect the last *security command mode message* used in AKA [13] and all subsequent ones, whereas we are able to provide integrity protection and encryption from the beginning. Mechanisms of delivering and activating smart cards (sending card, *Personal Identification Number* (PIN) and *Personal Unblocking Key* (PUK) per mail and activating it with the appropriate PIN) is also realized analogously in UMTS systems.

D. Public Key Infrastructure

Today, the exchange of public–keys is done via a public key infrastructure e.g., defined by the ITU-T standard X.509 [14][15]. All approaches have in common that a particular user or node publishes its public key on a key server of some sort from which it can be downloaded by other peers. After that, encrypted and signed messages can be exchanged. This, of course, requires that each participating node has to trust the key server. If a key pair ever gets lost, it can be revoked by including it in the so–called *Certificate Revocation List* (CRL), where all invalid keys and certificates are kept [15].

In [16] Ellison et al. argue that today's public key infrastructure based on *certificate authorities* (CA) imposes ten major risks. They describe, for example, the problematic trust background of self-proclaimed authorities and discuss the weakest link issue of the CA structure. Furthermore they raise the question how the certificate holder identifies himself against the CA. They state that several procedures do exist and that there are no consistencies over all CAs.

III. HiiMap Architecture

The HiiMap Next Generation Internet architecture [2] is based on the locator/identifier–split principle and provides a two–tier hierarchical mapping system. In the following, we will give a brief overview of the mapping system, as we will use HiiMap as example architecture.

In HiiMap, the mapping system is divided into so called regions as illustrated in Figure 1. Each region is responsible for all its nodes and has to provide the mapping for them. The region remains responsible, even if the node temporarily roams to another region. To identify a *responsible region* (RR) for a node, an additional 8 bit *regional prefix* (RP) to the identifier is provided. Within the HiiMap architecture, the identifier is assigned for life time and not subject to change as long as the owner doesn't request a new one. The regional prefix, however, is allowed to change whenever a node permanently migrates to another region. In HiiMap the identifier is called *unique identifier* (UID). The identifier and regional prefix is depicted in Figure 2.

Whenever a node wants to contact another node, it needs to query the RR of that node for the actual locator (which is called *local temporary address* or *LTA* in HiiMap). Therefore, it needs to know the regional prefix for that

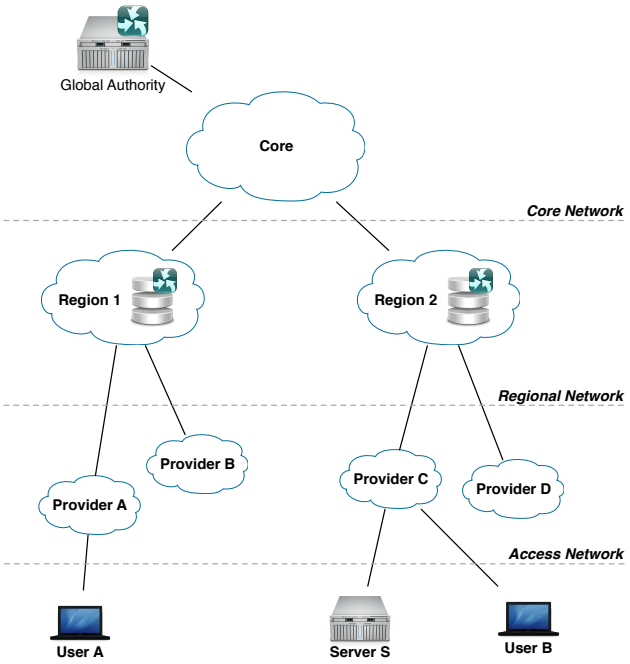


Figure 1. Example HiiMap topology

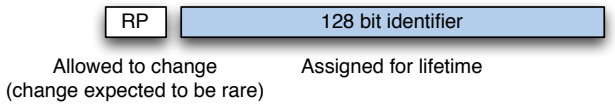


Figure 2. Identifier with regional prefix

region. In case the regional prefix hasn't been cached from previous communications, there are two ways to obtain it. The first possibility is together with the identifier itself. In case the identifier was learned from a link on a website or by means of a domain name system, the regional prefix can be provided along with the identifier. The second and fail proof possibility is to query the global authority. The global authority (GA) holds all $\langle RP, identifier \rangle$ tuples and can be queried, in case the regional prefix can't be learned by any other means.

As mentioned earlier, the mapping system is partitioned into multiple regions. For HiiMap, we propose to base the partitioning on countries, whereby each country forms its own region. This concept has two important advantages. Firstly, most countries show a relatively stable state. It rarely occurs that a country institutes or vanishes. This means there is a very seldom need to adjust the regional prefix. The second benefit of a partitioning based on countries is the common legal system. Each country has its own laws and ways of law enforcement. Therefore, a region based on more than one country has to deal with different political and legal systems. Smaller countries with similar laws, of course, can form a single region to lower the administrative overhead. In this way it is easier to build trust relationships between providers and handle infringements of contracts by the local law enforcement. Furthermore, we propose that the mapping system in each region is operated by a non–profit organization.

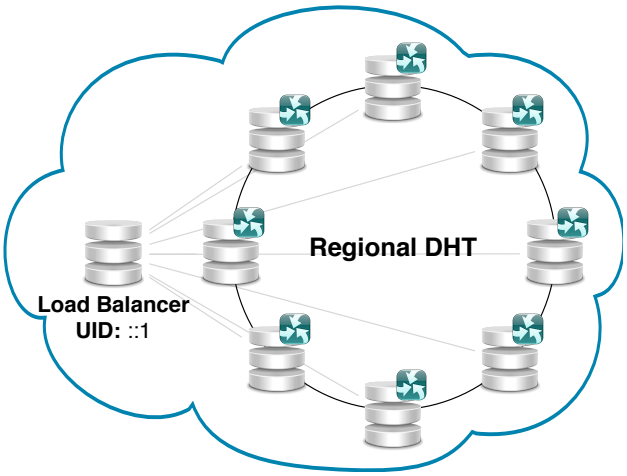


Figure 3. Mapping system of one region

Figure 3 illustrates the mapping architecture within one region. To be able to cope with the huge load of the mapping system, one–hop distributed hash tables (DHT) are used. In that way, it is no problem to meet the storage capacity requirements and the servers within the DHT are able to handle frequent locator updates and mapping requests. To provide a well–known address and fairly distribute the request load over the DHT, a load balancer is used. The address for each region is the same (e.g. region number::1) and clients do not need any additional information to access the mapping service of any region.

IV. Public Key Infrastructure

In this section, we describe the integration of the public key infrastructure into the HiiMap mapping system.

A. PKI and the mapping

In today's Internet, the public–key infrastructure is separated from all network services. This means that additional resources for the PKI must be provided despite all the network elements already in place for other functionality (e.g. DNS server). Contrary to this, we propose to integrate the PKI into the mapping system for the HiiMap architecture. This has the benefit that resources can be shared between functionalities and maintenance can be kept significantly lower compared to operating separate services.

Each mapping entry consist of the identifier as the primary key and a set of locators by which the node currently can be reached (see section III). Further, a timestamp of the last

update and a flag indicating whether the location update was cryptographically signed by the node or not is stored (we will come back to this issue later on).

To combine the PKI with the mapping system, only the public–key of each node must be additionally stored for each mapping entry. This means, that no additional protocol or infrastructure must be provided for querying and storing the public keys. Because the public–key is a very static value and not expected to change frequently, the additional burden for the mapping system is limited and the public–key databases can be optimized for frequent read requests—contrary to frequent read and write requests for the locators.

In comparison to today's public–key infrastructure, it is also not necessary to provide additional lists for key revocation. This is implicitly realized by the loose identifier - public–key binding. Whenever a public–key for a certain identifier changes, the old public–key implicitly becomes invalid.

B. Trusting the mapping

By storing the public–key at only one location (region) in the mapping system, however, the user heavily depends on the trustworthiness of that particular location. In case the mapping service provider collaborates with an attacker, it could send a wrong or manipulated public–key to the client. Therefore, any security functionality based on the public/private–key principle would be rendered useless. Even worse, the client considers the connection to be secure while in fact talking directly to the attacker.

Therefore, we propose to distribute several copies of the public–key to various independent locations (regions) in the mapping system. Figure 4 illustrates the basic principal.

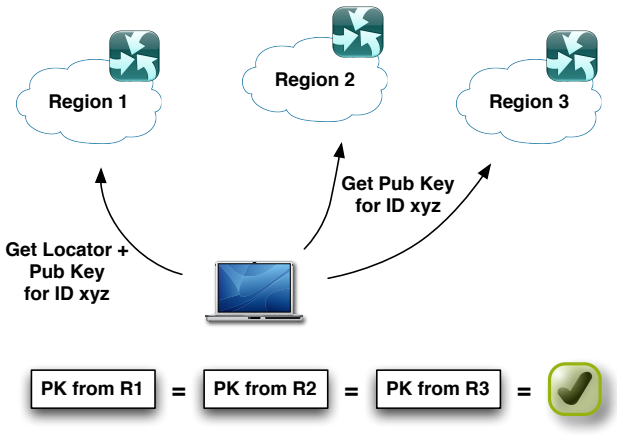


Figure 4. The public key is stored at multiple regions

The client first queries the responsible region (RR) of the identifier it wants to communicate with. As response, the RR replies with the locator and public–key stored for that identifier. In a next step, the client queries additional regions for the public–key. We will explain which regions to query in the next section. After receiving all requested public–keys, the client compares these. In case they do match, it is very likely that the public–key is the correct one. Contrary, if they differ, the client can either stop the communication setup or decide, which is the correct key based on the majority principle.

There is one special case, however. If the public–key from the RR differs from the other ones, then the retrieved locator must be considered incorrect as well. This is because having identified the RR as accessory or even the attacker itself, it is very likely that the locator has been modified as well and is now pointing directly towards the attacker.

A solution to this problem would be to also replicate the locator over several other regions. This, however, is not a good idea performance wise. The locator is the entry in the mapping system, which will be updated and changed frequently. In case several regions hold a copy of it, these changes have to be carried out to all of them. The public–key on the other hand is expected to change very rarely and thus causing very little update traffic.

C. Determining the storage location

Having copies of the public–key distributed over several locations in the mapping system, one questions remains: In which way does the client learn about the storage location of the additional copies.

Storing the list of the additional locations at the RR does not solve the problem. In case the RR is the attacker, it can simply manipulate this list as well and distribute the malicious key to collaborating regions. Therefore, the client must learn the information about the storage locations in another way.

For the following proposal, we assume that less than 256 regions exist and the key is distributed to two additional regions. After having received the locator and public key from the RR, the client hashes the identifier to an 16 bit value. The 16 bit value is split into two halves (8 bit each). Each 8 bit value represents the storage location of one of the public key copies. We will call it key storage address space (KSA) from now on. Since the 8 bit address space for the regions is not completely full, a mapping directive is required. This mapping directive can be downloaded from the global authority. It is sufficient to do this very seldom, as the directive is expected to change very rarely. For each value in the KSA, the mapping directive specifies a region, where the key is stored. This means, that a single region can be responsible for several KSA values. In that way, the load can be fairly distributed over all regions depending on their size. Figure 5 illustrates the process.

Should the hashing and mapping to regions result in two copies of a key being stored at the same region, the first copy is stored at this region and the second copy at the region with the next higher region number.

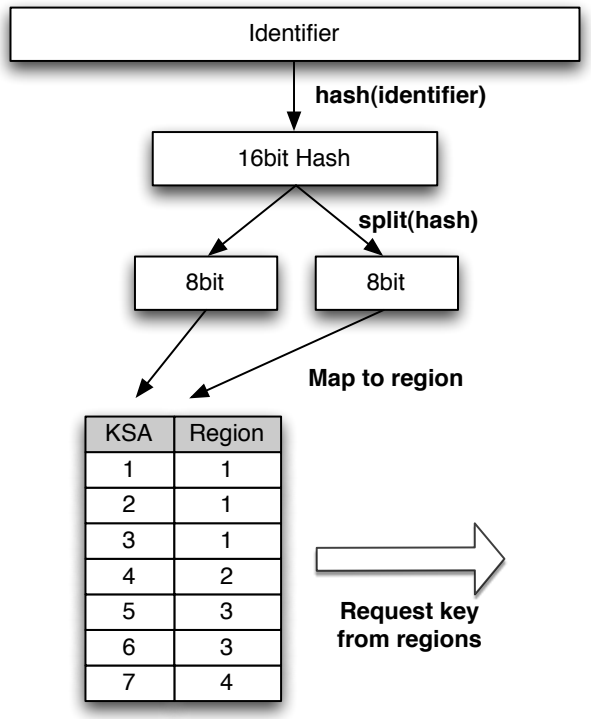


Figure 5. Retrieving the additional key storage locations

V. User Key Management

The user key management in our concept is based on cryptographic smart cards (herein after referred to as *smart card*). For the sake of simplicity, we assume that the smart card can't be compromised on a physical level and that it is protected by the well–known PIN/PUK mechanism. As of today, this mechanism is considered to be secure enough to protect any smart card from unauthorized access.

As with the previous section, the approach is not bound to a specific architecture. We again, however, will use HiiMap as an example architecture to illustrate the functionality. Any other concept, which provides a *single point of trust* (SPT) can be used as underlying architecture. As with HiiMap, the SPT should be independent—politically and organizationally—to reflect all participants' needs and interests in the same way. It will take responsibility considering key management issues and authentication of particular nodes later on. SPT, however, doesn't mean that it has to be one physical component with respect to reliability. Furthermore, management tasks can be delegated to subsequent authorities. In HiiMap the global authority (GA) acts like a SPT and can delegate tasks to the regional authorities (RA).

In this section we will discuss the topics of authentication methods, initial bootstrap, key revocation and how devices with low computational power can participate.

A. Peer communication

The authentication and communication concept differentiates between the used hardware components (notebook, PDA, mobile phone, etc.) and access authorization, which is handled by the above mentioned smart card. Vendors only have to provide an interface for this card in each of their products. The assembly of these smart cards is done by the *single point of trust* (SPT) respectively by another party authorized and trusted by the SPT. They contain a master key pair, the card-ID and, of course, the identifier address. In HiiMap, this identifier is called UID (see section III). Before the SPT can send the cards to authorized providers, it has to save every public–key stored on them and the other entries already mentioned in Section IV-A. The SPT furthermore saves the appropriate card-ID and if the particular card is already in use or not. Therefore, it always possesses all relevant information. We have to remark that authorized providers of course can keep a certain amount of smart cards in stock so that they do not have to request every single card each time they get a new customer.

Every time somebody buys a new device, he chooses a provider. If he is in possession of such a smart card already, he can either sign up for a new one (and meanwhile use the existing card) or use the old one in the new device. It is also possible to change the provider with every card request. This modularity is an important advantage of smart cards in comparison to fixed security modules as they provide much more flexibility and do not involve manufacturers in the network management process (assignment of identifiers, etc).

If the user requests a new card, the provider then sends smart card and PIN/PUK to the user, for example by mail. If the user requests such a card for the first time, the trader informs the particular provider directly at buying time to minimize downtime. At the same time, the provider tells the authorities (the SPT or its delegates) about the selling of this card. They can then update their databases and know that the particular card is in use from now on. Thus, the authorities know which cards are in use and which aren't at any point in time. This makes it difficult for possible attackers to use non–assigned card-IDs. Assigned IDs cannot be compromised, because the attacker cannot prove the possession of the private key, as we will see later on.

After the user has received the smart card, he can sign on to the device by inserting the card and typing the correct PIN. The security mechanisms can then be enabled and the device is able to authorize itself to the network.

B. Bootstrap

The procedure of joining a network for the first time is called *bootstrap*. If the particular user is not yet known to the network and other users, there is no possibility to prove his identity in general. In most cases of security mechanisms, other peers have to trust this user once. After keys have been

exchanged between participants, they can later on check the identity with the corresponding key pairs. As this is a great drawback (possible attackers could replace the keys with their own), we will present a solution to this problem.

As mentioned above, the authentication procedure uses a smart card for key storing and cryptographic functions. The SPT (and its delegates—also called "authorities" in the following) is in possession of all public keys stored on these cards and can connect them to the respective IDs (see section V-A). This is the essential point of the bootstrap mechanism. Imagine a node i joining the network where the associated user has already enabled the smart card by entering the correct PIN. As shown in Figure 6, the node first has to send a *location update request* to the responsible authorities. It contains the card ID of the smart card used, so that the network can check whether or not the card is allowed to participate in the network's functionalities. This message is already signed with its private key to prove integrity. Therefore, the whole communication is integrity protected from the beginning.

The authority can then lookup the node's public key. If the public key is not yet known to the authority, it has to request it from the SPT. After that, the authority computes a common secret K_{ir} using node i's public and its own private key, similar to the Diffie-Hellman–procedure [7]. This secret K_{ir} can also be computed by node i in the same way (i also gets the public key of authority R from the SPT). Therefore, the common secret K_{ir} never has to be exchanged between the two peers, which eliminates the danger of being compromised. Furthermore, it is only used once to encrypt data (part of message 2 in Figure 6). With this message, the authority chooses a random session key K_p and a rule to generate a modified common key K_{ir}^*. K_{ir}^* can be calculated, for example, by shifting K_{ir}, computing the product K_{ir} XOR itself or other methods. The authority can then answer the location update request by sending this message containing the rules for generating K_{ir}^* and the security functions the authority is capable of (message 2 in Figure 6). This information is encrypted using the random session key K_p. K_p is encrypted with the common secret K_{ir} and sent inline in the packet (message 2 in Figure 6), based on the principle used in SKIP [17]. Besides, K_{ir} is only used once to encrypt data. All other packets use the modified version, which again minimizes the risk of compromising K_{ir} itself. The header information in this and all other subsequent packets are sent in plain text. This reduces complexity for network nodes, firewalls and so on. A possible attacker possibly acquires part of the payload by resolving the security functions by sending an own location update request. However, this is not enough information to decrypt the key. Afterwards, the authority sends another packet to node i containing a random number n_A, which is again encrypted using K_p (and K_p with the *modified* common secret K_{ir}^*), see message 3. Node i can extract the chosen session key K_p by decrypting it with the common secret K_{ir} of message 2 and then the security functions and rules for generating K_{ir}^* with K_p. After that, node i can compute the modified common secret and therefore decrypt the random number n_A of message 3. Node i

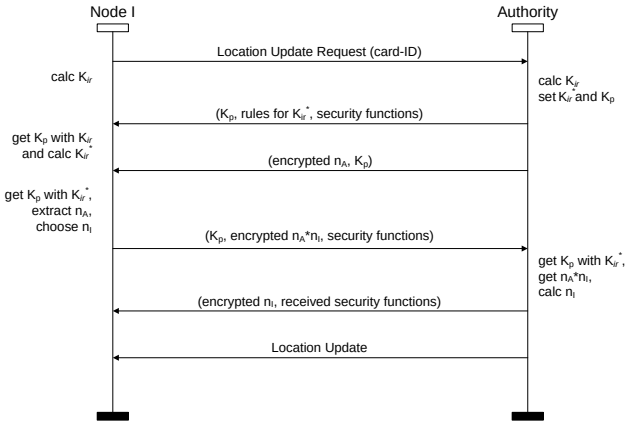


Figure 6. Bootstrap Message Flow Chart

then also chooses a random number n_I and calculates the product $n_A * n_I$. This product is sent back to the authority together with i's security functions. Both are again encrypted using a random session key K_p, which can be the same as above or vary depending on whether the key is valid packet- or session-wide. K_p again is encrypted using the modified common secret K_{ir}^* (message 4). The authority can extract the particular information in the same way node i did before and therefore calculate the chosen random number n_I. Afterwards, the authority can select the strongest algorithms for creating new session keys K_p supported by i and R. By sending back this random number in message 5 of Figure 6 both parties can be sure that the other part is in possession of the right (modified) keys. Additionally, the received security functions of node i are sent back to prove they have not been manipulated. A modification of all those messages would also mean that the signatures become invalid as every message is not only encrypted, but also signed to prove integrity. In a last step, the location update request is accepted by the authority, which results in publishing node i's assigned locator address in the mapping system, so that other nodes can resolve it from then on. The user or node is then allowed to upload his own key pair for further use, which has to be validated with the old key pair again. Thereby, the peer has flexibility to use own algorithms for creating the keys and the possibility to influence the parameters, such as the key length. If keys are changed later on, every party can signalize the wish to update it with a special *key_update* message. It contains the new key validated with the old one. Bootstrap is completed and node i can go on communicating with other peers.

Erroneous messages are ignored by the system and the

user has to resend them. To avoid denial of service attacks by exploiting this, the system only allows a maximum number of requests and responses at a time, i.e., five requests within ten seconds. After that, the system will not accept any more messages of the particular node in a certain time. At any time, the user has to be informed if encryption is disabled. This can be done by the operating system, for example.

C. Dealing with network components

Cases may occur, where users are located behind a firewall, proxy or similar network entities and are not directly reachable. These non–end–to–end cases are considered as well. A typical connection establishment for those cases looks like the following:

First of all, the requesting peer contacts the firewall by looking up the appropriate UID associated with a human readable address as known from the *Domain Name System* (DNS). The firewall redirects the connection request to the particular node, i.e., according to load balancing rules and also informs the requesting peer (see message 2 in Figure 7). Of course, UIDs have to be looked up by the firewall, too. We call them *RSLVreq (resolve request)* and *RSLVresp (resolve response)*, but left them out in Figure 7 for the sake of simplicity. After that, both nodes can request the

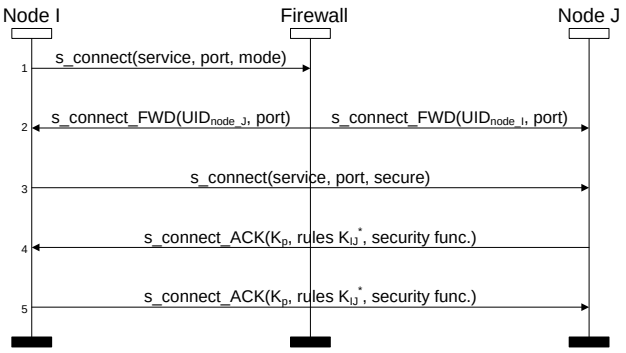


Figure 7. Passing firewalls (simplified)

needed information held by the SPT and start connecting to each other (see messages 3 to 5 in Figure 7, where some additional parameters for encryption and defining the desired service are negotiated—detailed explanation of single parameters see Section V-B). The firewall itself is able to route the packets correctly as it stores the connection data like in NAT–gateways. By using clear headers instead of encrypted ones, every network node and therefore the firewall is able to process all needed data. Both users can decide about the connection mode on their own at any time, e.g., encrypted or plain. Similar to the bootstrap process (see Section V-B), keys can be updated anytime by sending a *key_update* request.

D. Disabling authentication

In some cases it may be necessary to connect even devices without smart cards, as they are difficult to reach physically, e.g., sensors, satellites, etc. Moreover, most of the available sensor data is not crucial, so that there is no drawback to operate them in plain text communication mode. Additionally, not every single sensor needs to be connected to the Internet, e.g., in cars it is sufficient if the board computer is connected. Nevertheless, cases may occur in which those devices have to be integrated without the chance of attaching the smart card to them. The procedure then is as follows: First of all, we assume a legal owner of this device, let us say, a company operating a sensor. This owner requests a smart card for the sensor in the described manner. After that, he securely keeps the card somewhere and implements the particular UID into the sensor's firmware and also his own UID. Concurrently, the owner connects to the network using the sensor's smart card and the appropriate PIN. After the encrypted location update is completed successfully, he then tells the network or alternatively the authorities that the UID he is connecting from will disable authentication mode in the future. This is stored in the database entry called *mode of last location update* (see Section VI-B). The network then and only then permits plain text location updates from the particular UID. Thus, disabling security functions is possible, but only on explicit inquiry. If the UID ever wants to return to secure mode, this again has to be done with the appropriate smart card and PIN and can therefore only be realized after secure location updates. After that, the owner keeps the card and PIN secret again. In this way it can always be guaranteed that only the sensor itself can disable encryption. To summarize, the procedure is depicted in Figure 8.

Consequently, an attacker is not able to force plain text communication. If the owner decides to sell the device, he simply has to distribute the particular card to the new owner, whereon he can handle communication modes on his own. Requesting a new smart card for the sensor (with same ID, of course) is also possible. This mechanism is another great advantage of modular chips in comparison to fixed ones as they easily enable such devices to join the network. Every time such a sensor connects to the network, it sends an unencrypted location update request to the responsible authorities. The mode of the connection (secure or plain) is indicated by special fields in the message's header. The authorities can then check, whether plain text location updates are enabled for the particular node or not. After that, the location update without authentication is granted and the sensor can participate in the network's functionalities. We have to remark that such devices only get limited access to resources as they have not been identified securely. Furthermore, they have to keep plain text communication enabled for security reasons (otherwise one could easily

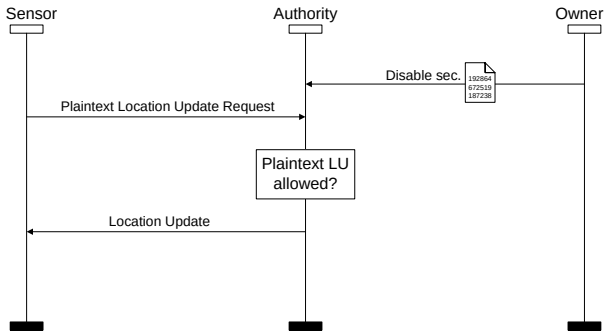


Figure 8. Location update with disabled security

upgrade to ”secure” mode and get access to critical content).

Every time a user requests the public key of another user from the authorities, they also inform him about the last kind of location update (secure or insecure). He can then decide on his own whether he wants to continue connecting to the desired node or not. This offers many possibilities for user-defined security policies, as it is possible, e.g., to grant access to e-business-products only to securely authorized nodes, whereas insecurely authorized nodes only get access to information material.

The concept of providing plain text authentication can also be seen as fallback solution in case the whole system becomes compromised in the future. Then it is sufficient to switch all participants to plain text mode and provide an overlay network that is responsible for security issues.

E. Key Revocation

The high modularity in this concept using smart cards also implies a drawback: the card can easily be lost or even stolen. This can be mitigated as the card itself is not usable without the appropriate PIN assigned to the authorized user. In our opinion, the chance for a possible attacker to get the valid PIN in only three guesses is very small and thus negligible. Besides, after that the card is disabled until the correct PUK (*personal unblocking key*) is entered. The PUK again may not be entered incorrectly more than ten times. This behavior is known from today's concepts and is assumed to be secure enough. The smart card itself—as in today's ones—has to be protected against physical and chemical manipulation such as side channel attacks, power analysis, etching and so on (details see [18]). Baring these things in mind, the physical theft of a smart card implies no great security risk. Of course, there has to be some kind of approach the user has to follow if his card is stolen or lost:

The user has to report the loss or theft immediately to his responsible provider, i.e. by phone, who then disables the card by denying the particular card-ID from joining the network in the future. This can be done by the authorities or delegated to the providers. The disabled card-IDs are stored in order to detect future connections of the card and eventually having the chance of locating it. The SPT then also has to be informed as it now is in charge of producing a new card containing the old identifier (as the identifier shall not change), a new key pair and, of course, a new card-ID. This step may take some time, as the card is not already produced but has to be created on inquiry. After that, the normal procedure takes place again: The SPT and the responsible authorities get the public key and replace the old one with it. The card/PIN pair is delivered to the provider who then ships it to the desired customer. After that, the user can proceed as normal.

The provider has to make sure that only the actual owner of the card is allowed to report the loss or theft so that an attack on disabling all cards by simply calling all providers is not possible. This can be done by requesting some additional information of the caller, for example street, postal code, birth date etc. Even if the attacker is in possession of this information, the risk of such an attack is highly improbable as this causes much effort for the attacker and can not be automated in an easy way.

Following this procedure, the device or the user is able to request a new key pair without loosing his assigned identifier. Other nodes in the network will probably not even notice the change as they request the public key from the authorities and do not cache the old public key for an unlimited time.

VI. Evaluation

The security mechanism proposed in this paper is very flexible and neither bound to a specific algorithm nor architecture. Only some already mentioned pre-requirements must be met. For the analysis of the mechanism, however, we need to assume some protocols and algorithms, which are likely to be used. Please note that the mechanism can also be applied to different proposals and is not limited to the ones discussed in this section.

A. Algorithm

As already mentioned, we choose HiiMap as the underlying locator/identifier–split architecture where the *Global Authority* (GA) acts as single point of trust. The GA can delegate management and maintenance tasks to *Regional Authorities* (RA), which are responsible for their region respectively. Thus, without any security capabilities enabled, the system already has to store these entries (UID, LTA and Region Prefix).

The public/private–key principle requires an asymmetric cryptographic algorithm. The probably best known one is *RSA* by Rivest et al. [19]. RSA is widely used in electronic commerce protocols, and is believed to be secure given sufficiently long keys. While a key length of 1024 bit is still assumed to be secure enough, a length of 2048 bit is recommended (as of early 2010). A downside of the

RSA algorithm, however, is its large memory footprint and requirement towards the computational power of a device.

Another asymmetric algorithm, the *ECC* by Koblitz et al. [20], can be found on some low power units like smart cards and sensors. This is because ECC is less computational power and memory consuming compared to RSA [21]. Also the required key length is almost one magnitude smaller while maintaining the same level of security. An 160 bit ECC key, for example, is believed to be equal in terms of cryptographic strength compared to a 1024 bit RSA key.

To provide a better overview and point out the flexibility once again, we will calculate the resource requirements for both, RSA and ECC. As you will see, replacing the algorithm will result in a huge decrease of requirements. In case of ECC, we choose a key length of 160 bit. We will first do the calculation with the RSA algorithm and then present the respective values for ECC.

B. Analyses

To calculate the overall overhead and storage requirements, we have to make some assumptions and declarations. First of all, we have to choose an average key length for the *public–key*. The initial key pairs stored on the smart cards are all of the same size, but users can later change them and influence these parameters. As in today's *Trusted Platform Module* (TPM) [22], which we consider to be safe enough, RSA keys of 2048 bit are used, we will assume an average *public–key length* of 2048 bit for the RSA study. In fact, this is the entry which consumes most of the needed storage in comparison to the rest. Increasing or decreasing this value will have a strong impact on the overall storage capacities. The next entry is the smart card's *actual allowed ID*. With this ID, the network can verify if the particular node is allowed to join the network or if this card has been reported as stolen or lost. The card-ID is only valid within the UID range, therefore it is sufficient to reserve 32 bits for it. This value is high enough to avoid guessing the next valid card ID by attackers, as well as leaving sufficient space to replace the card several times every day. The last entry in Table I *(list of disabled card IDs)* is the opposite: it holds all card IDs that have been disabled and are no longer allowed to connect. In case of such an ID joining the network, the authorities can trace the request and thereby locate the missing card. Depending on the number n of disabled smart cards, the list may increase.

Last but not least, there are two entries of four bit each: On the one hand, the field *UID assigned*, which specifies if the particular UID is already in use by a node or not, and on the other hand, the entry *mode of last location update*. This field gives information about whether the last location update was encrypted or in plain text. It is sent with every public key request so that the peer can decide on its own whether or not it wants to continue connecting. Both fields could have also been realized with only one bit, but by reserving three additional bits we get enough flexibility to adapt future challenges, e.g., the connection modes can be split up in more detail.

entry	length
UID	128 bit
LTA	128 bit
RP	8 bit
public key	2048 bit
valid card ID	32 bit
mode of last location update	4 bit
UID assigned	4 bit
list of disabled card IDs	$n*$ 32 bit
Sum	$2352 + n * 32$ bit

Table I
ENTRIES TO BE STORED (LIKE IN HIIMAP [2]) – RSA

A summarizing overview of all necessary entries can be found in Table I. Based on them we want to present a typical example to estimate the needed storage capacities in a future NGI system. Therefore, we choose ten as the number of smart cards per human being on earth (currently about $6.7 * 10^9$). We assume an average invalid card–ID count of 20 per smart card. Putting all these parameters (see Table II) together, we get a total requirement of $10*6.7*10^9*(2352+20*32)$ bit $= 2.00464 * 10^{14}$ bit $=$ **25.05 Terabyte**. Even in today's architectures, this value is no major challenge. Taking the computing power available in 10 to 15 years into account, we are not talking about a huge burden compared to the security benefits we are gaining. If the load can be delegated to subsequent authorities like the RAs in HiiMap, the burden is distributed over several nodes. This also applies to bandwidth and other metrics so that the SPT itself ideally is not involved in handling authorization requests unless the delegates (RAs in HiiMap) can not resolve them on their own. Thus, our concept demands no great resources and therefore is suited for use in any kind of locator/identifier–split architecture.

parameter	value
n	20
cards per human being	10
human beings on earth	$6.7 * 10^9$

Table II
PARAMETERS USED (LIKE IN HIIMAP [2])

Having outlined in detail the requirements for an implementation with RSA, we will now shortly present the respective values in case of using ECC. Most of the values presented in Table I do not change as they do not depend on a specific algorithm. What in fact does change, is the public–key, of course. This decreases from 2048 bit (RSA) to 160 bit (ECC). So the overall sum decreases, too (see Table III). The values presented in Table II do not change

entry	length
UID	128 bit
LTA	128 bit
RP	8 bit
public key	160 bit
valid card ID	32 bit
mode of last location update	4 bit
UID assigned	4 bit
list of disabled card IDs	$n*$ 32 bit
Sum	$464 + n * 32$ bit

Table III
ENTRIES TO BE STORED (LIKE IN HIIMAP [2]) – ECC

at all. If we calculate the sum again, we get an overall storage requirement of $10 * 6.7 * 10^9 * (464 + 20 * 32)$ bit $= 1.10 * 10^{13}$ bit = **1.38 Terabyte**. So the storage capacity can be decreased by an order of more than ten.

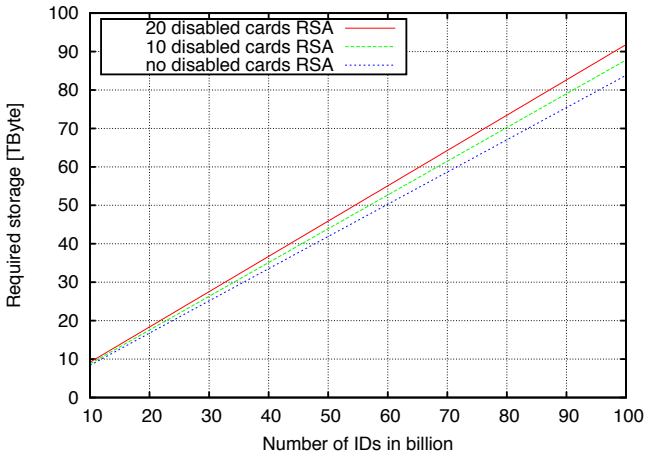


Figure 9. Storage requirement RSA

As already pointed out in section IV-A, we need to store the public key not only at one, but at multiple locations. For the overall storage calculation we have to reflect this additional requirement as well. The total storage varies depending on the number of used UIDs and the number of disabled card IDs per UID. Figure 9 shows the storage requirement using 2048 bit RSA keys and Figure 10 the requirements with 160 bit ECC.

Concluding our computations, the concept does not consume a huge amount of resources. No matter which algorithm we choose, we are not facing a huge burden to the architecture—even in today's view. Nevertheless, by decreasing the biggest factor, the public–key length, we can decrease storage requirements drastically.

VII. CONCLUSION

Many security concepts for locator/identifier–split architectures bind the identifier to the public–key. Contrary to this common approach, we suggested a loose coupling between

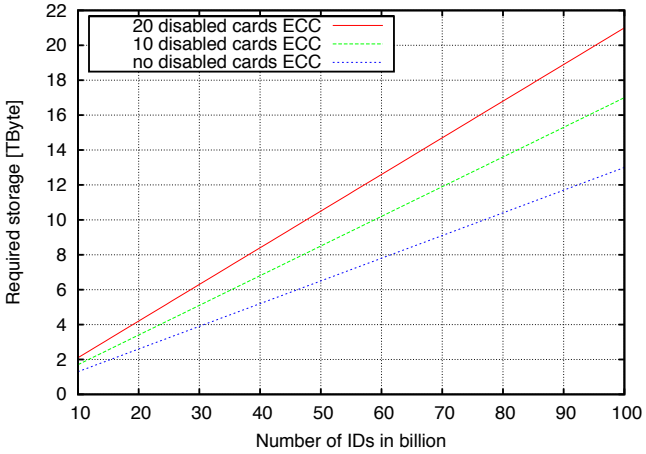


Figure 10. Storage requirement ECC

the identifier and the public–key, allowing for exchangeability of those two entities [2].

In this article, we presented a holistic approach to public key management including key distribution, key revocation and key storing. We covered the public–key infrastructure aspect by extending the mapping system to also store a public–key for each identifier. We introduced a mechanism to trustfully retrieve keys from the mapping system without being dependent on a single region of the mapping.

Furthermore, we discussed the client side key management and suggested to use smart cards to store the private and public–key. We described the initial bootstrap process, detailed the communication setup and showed how devices with very limited computational power can also participate by disabling encryption.

The concept is not bound to a specific crypto algorithm and is able to cope with varying key length. Therefore, the architecture is very flexible and open to future improvements or requirements. Although we explained the concept by using the HiiMap architecture as example, the concept can be applied to any locator/identifier–split architecture, which provides a single point of trust and a mapping, which can be divided into several administrative zones.

ACKNOWLEDGMENT

This work has been performed within the G-Lab project and was funded by the Federal Ministry of Education and Research of the Federal Republic of Germany (Project ID 01BK0807). The authors would also like to thank their colleagues at the Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities (see http://www.lrz.de/) for helpful discussions and valuable comments about this paper. The authors alone are responsible for the content of the paper.

REFERENCES

[1] W. Fritz and O. Hanka, "Smart Card Based Security in Locator/Identifier-Split Architectures," *International Conference on Networking*, pp. 194–200, April 2010.

[2] O. Hanka, G. Kunzmann, C. Spleiß, J. Eberspächer, and A. Bauer, "*HiiMap: Hierarchical Internet Mapping Architecture*," *In First International Conference on Future Information Networks, Beijing, China, P.R. China*, pp. 17–24, October 2009.

[3] T. Koponen, M. Chawla, B.-G. Chun, A. Ermolinskiy, K. H. Kim, S. Shenker, and I. Stoica, "A data-oriented (and beyond) network architecture," in *SIGCOMM '07: Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications*. New York, NY, USA: ACM, 2007, pp. 181–192.

[4] R. Moskowitz and P. Nikander, "Host Identity Protocol," IETF, United States, RFC 4423, May 2006.

[5] M. Menth, M. Hartmann, and M. Hoefling, "Firms: a future internet mapping system," *IEEE Journal on Selected Areas in Communications (JSAC), Special Issue on Internet Routing Scalability*, August 2010.

[6] A. Feldmann, L. Cittadini, W. Mühlbauer, R. Bush, and O. Maennel, "HAIR: Hierarchical Architecture for Internet Routing," in *ReArch09*. New York, NY, USA: ACM, December 2009.

[7] W. Diffie and M. Hellman, "New Directions in Cryptography," *IEEE Transactions on Information Theory*, vol. IT-22, no. 6, pp. 644–654, 1976.

[8] M. Komu and J. Lindqvist, "Leap-of-faith security is enough for ip mobility," in *Consumer Communications and Networking Conference, 2009. CCNC 2009. 6th IEEE*, Januar 2009, pp. 1–5.

[9] S. Siddique and M. Amir, "GSM Security Issues and Challenges," in *Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2006. SNPD 2006. Seventh ACIS International Conference on*, June 2006, pp. 413–418.

[10] M. Toorani and A. Beheshti Shirazi, "Solutions to the GSM Security Weaknesses," in *Next Generation Mobile Applications, Services and Technologies, 2008. NGMAST '08. The Second International Conference on*, September 2008, pp. 576–581.

[11] M. Khan, A. Ahmed, and A. Cheema, "Vulnerabilities of UMTS Access Domain Security Architecture," in *Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2008. SNPD '08. Ninth ACIS International Conference on*, August 2008, pp. 350–355.

[12] A. Bais, W. Penzhorn, and P. Palensky, "Evaluation of UMTS security architecture and services," in *Industrial Informatics, 2006 IEEE International Conference on*, August 2006, pp. 570–575.

[13] U. Meyer and S. Wetzel, "A man-in-the-middle attack on UMTS," in *WiSe '04: Proceedings of the 3rd ACM workshop on Wireless security*. New York, NY, USA: ACM, 2004, pp. 90–97.

[14] R. Housley, W. Ford, W. Polk, and D. Solo, "Internet x.509 public key infrastructure certificate and crl profile," IETF, United States, RFC 2459, January 1999.

[15] R. Housley, W. Polk, W. Ford, and D. Solo, "Internet x.509 public key infrastructure certificate and certificate revocation list (crl) profile," IETF, United States, RFC 3280, April 2002.

[16] C. Ellsion and B. Schneider, "Ten risks of pki: What you're not being told about public key infrastructure," *Computer Security Journal*, vol. 16, no. 1, pp. 1–7, 2000.

[17] A. Aziz, M. Patterson, and G. Baehr, "Simple Key-Management for Internet Protocol (SKIP)," in *Internet Society: INET'95 Hypermedia Conference Proceedings*, June 1995.

[18] W. Rankl and W. Effing, *Smart Card Handbook*, 3rd ed. John Wiley & Sons, 2003.

[19] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Commun. ACM*, vol. 26, no. 1, pp. 96–99, 1983.

[20] N. Koblitz, "Elliptic curve cryptosystems," *Mathematics of Computation*, vol. 48, no. 177, pp. 203–209, 1987. [Online]. Available: http://www.jstor.org/stable/2007884

[21] N. Gura, A. Patel, A. Wander, H. Eberle, and S. C. Shantz, "Comparing Elliptic Curve Cryptography and RSA on 8-bit CPUs," in *Lecture Notes in Computer Science*. Berlin/Heidelberg, Germany: Springer, 2004, vol. 3156/2004, pp. 925–943.

[22] The Trusted Computing Group, "Trusted Platform Module (TPM) Main Specification, Version 1.2, Revision 103," http://www.trustedcomputinggroup.org/resources/tpm_main_specification,17.11.2009.

Security Capacity of the Fuzzy Fingerprint Vault

Johannes Merkle, Matthias Niesing, Michael Schwaiger
secunet Security Networks AG
D-45128 Essen, Germany
johannes.merkle@secunet.com,
matthias.niesing@secunet.com
michael.schwaiger@secunet.com

Heinrich Ihmor, Ulrike Korte
Bundesamt für Sicherheit in der Informationstechnik
D-53175 Bonn, Germany
heinrich.ihmor@bsi.bund.de
ulrike.korte@bsi.bund.de

***Abstract*—We investigate the security of a privacy enhancing technique for fingerprint authentication known as *fuzzy fingerprint vault*. This technique is based on the *fuzzy vault* of Jules and Sudan, a scheme that allows error tolerant authentication, while preserving the privacy of the reference data. We explore if and under what circumstances a secure fuzzy fingerprint vault can be implemented. We derive both upper and lower security bounds for any attacks that attempt to recover the template from the stored reference data, and, at the same time, significantly improve the best known attack. Furthermore, we show how to select optimal parameters and evaluate both minimum minutiae match rates and minimum number of minutiae needed to obtain an appropriate security level. Our results quantify the security capacity of the fuzzy fingerprint vault and provide important tools for selection of suitable parameters.**

***Keywords*-biometric template protection; fingerprint; fuzzy vault; polynomial reconstruction**

I. Introduction

Without any doubt, fingerprints are the biometric traits most widely deployed for authentication. However, the storage of biometric reference data introduces considerable risks for biometric authentication systems and raises serious concerns regarding privacy and data protection. One of the most prominent solutions to solve this issue is the fuzzy fingerprint vault, which allows error tolerant fingerprint authentication while preserving the privacy of the biometric features [1]. It belongs to the class of *biometric template protection* techniques [2], and is based on the fuzzy vault scheme [3] of Juels and Sudan, which applies Reed-Solomon decoding to redundantly bind the biometric template to a randomly selected secret polynomial.

Fingerprint authentication is typically based on minutiae, which are specific features of the fingerprint pattern. The variety and extent of errors in minutiae measurements, particularly, frequent insertions, omissions and re-ordering of the measured minutiae, pose a considerable challenge to template protection schemes [4]. The fuzzy vault is able to tolerate such errors and, hence, is particularly interesting for minutiae-based authentication.

Several publications [5][6][7][8][9][10] report successful implementation of the fuzzy vault scheme based on minutiae. However, the subsequent publication of efficient attacks [11][12] demonstrates that the parameters proposed do not provide adequate security.

For the fuzzy vault, theoretical results are known, from which rigid security estimates could be deduced. In particular, Dodis, et al. [13] proved upper bounds for the information leakage by the stored data, which determines the maximum success probability of an attack trying to guess the template or the key from the stored reference data, see Section IV-B for details. In addition, an attacker's success probability depends on the original entropy of the biometric feature vector - or, equivalently, its redundancy. Therefore, a realistic estimation of the entropy of the biometric feature vector is a key aspect for a sound security analysis.

On the other hand, these provable lower security bounds are not sharp. Firstly, these bounds only estimate the success probability of attacks and do not consider the effort required for each trial. Secondly, the proof techniques used in [13] overestimate the information leakage. Achieving provable security may be a very appealing objective, but it is also interesting to determine how secure the scheme is in practice.

In this publication, we explore if and under what circumstances a fuzzy fingerprint vault can be secure with respect to both provable security and real attacks. In particular, we generalize the bounds of [13] to the case where the minutiae and chaff points are chosen with a minimum distance to reduce false matchings, and also give an exact estimate for the entropy of a feature vector consisting of minutiae location data. On the other hand, we estimate the effort required for practical attack methods and present an improvement of the best known attack. Then, we show, how the parameters can be optimized and determine minimum minutiae match rates with respect to both provable security and practical security.

This article is structured as follows. In Section III, we give a description of the scheme. In Section IV, we conduct a theoretical analysis of its security and error robustness both with respect to information theoretical results and practical attacks. Section V presents methods for parameter optimization with respect to the deduced security bounds, and Section VI provides results using empirical data. A conclusion is given in Section VII.

II. Background

The fuzzy fingerprint vault is one of many template protection techniques that have been proposed in the literature, for instance, the *Biometric Encryption* scheme by Soutar et. al. [14], *Cancelable Biometrics* by Ratha et. al. [15], robust bit extraction schemes based on quantization, e.g. of Linnartz and Tuyls [16], of Chang et. al. [17], and of Chen et. al. [18], and applications of the fuzzy commitment scheme of Juels and Wattenberg [19] to biometric templates, e.g., the constructions of Martini and Beinlich [20] for fingerprints, of Kevenaar et. al. [21] for face recognition, of Hao et. al. [22] for iris, and of Korte et. al. [23] for DNA fingerprints. The fuzzy vault has also been applied to iris recognition, e.g., in [24].

A. The general fuzzy vault scheme

The fuzzy vault has been proposed by Juels and Sudan in [3] and [25]. It is an error tolerant authentication scheme based on the set of private attributes $m_1, \ldots, m_t$, e.g., biometric feature data. While the reference data stored (the *vault*) allows performing the authentication check, it does not reveal these attributes. The scheme deploys a variant of Reed-Solomon decoding and hides the private user data among a large number of random *chaff points*.

During enrollment of a user, her (pairwise distinct) private attributes are encoded as elements $x_1, \ldots, x_t$ of a finite field $\mathbf{F}_q$. Then a random secret polynomial $P(z)$ over $\mathbf{F}_q$ with degree smaller than k is chosen. Each of the encoded attributes x_i is evaluated over the polynomial, resulting in a list of pairs $(x_i, y_i) \in \mathbf{F}_q^2$ with $y_i = P(x_i)$. In order to hide the private attributes, $r - t$ *chaff points* $x_{t+1}, \ldots, x_r \in \mathbf{F}_q$ are randomly selected so that $x_i \neq x_j$ for all $1 \leq i < j \leq r$. For each chaff point x_i, a random $y_i \in \mathbf{F}_q$ with $y_i \neq P(x_i)$ is chosen. The list of all pairs $(x_1, y_1), \ldots, (x_r, y_r)$, sorted in a predetermined order to conceal, which points are genuine and which are the chaff points, is stored as the *vault*.

The redundant encoding of the polynomial using the genuine points and its hiding among the chaff points is illustrated in Figure 1.

For authentication and recovery of the secret polynomial, another set of attributes (the *query set*) has to be presented. This set is compared with the stored

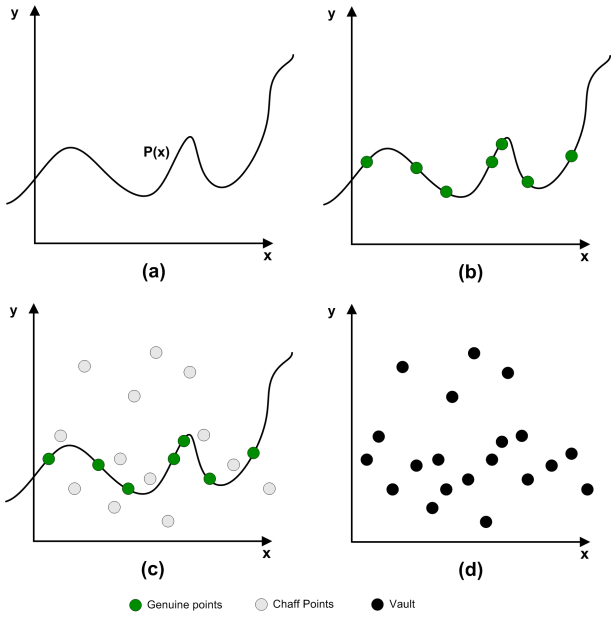


Figure 1. Illustration of the redundant encoding of the polynomial using the genuine points ((a) and (b)) and its hiding among the chaff points ((c) and (d)).

fuzzy vault $(x_1, y_1), \ldots, (x_r, y_r)$, and those pairs (x_i, y_i) are selected, for which x_i corresponds to an attribute in the query set. The selected points are then used to try to recover the secret polynomial using Reed-Solomon decoding.

If the number of genuine points among the identified correspondences (*correct matches*) is at least k, the secret polynomial can be recovered, either by Reed-Solomon decoding or by polynomial interpolation. However, if the set of correspondences also comprises chaff points (*false matches*), the number of correct matches must be greater than k, or the decoding must operate on subsets of the matches resulting in many trials. Details are given in Section III-B3.

B. Previous results

In [3], Juels and Sudan already provided an information theoretical security analysis for the general fuzzy vault scheme by giving estimates for the number of candidate polynomials that would fit with a given vault. A comprehensive information theoretic treatment of the fuzzy vault was given by Dodis et. al. [26][13], who proved upper bounds for the loss of entropy (information leakage) by the stored data for the fuzzy vault, the fuzzy commitment, and other schemes. In [1], we applied these general results to the fuzzy fingerprint vault and deduced lower bounds for the number of required minutiae and minutiae match rates, i.e., the fraction of minutiae in the vault matching with the minutiae of the query fingerprint, were deduced.

Implementations of the fuzzy vault for fingerprints were reported in many publications, the most notably of which are described below.

Clancy et. al. [5] were the first to propose a fuzzy fingerprint vault. Their construction uses only the location information of the minutiae, i.e., their orientations are neglected, and uses several measurements of the minutiae during enrollment to filter out spurious or unreliable minutiae. A drawback of their implementation is that it assumes that the fingerprints are already prealigned. The security against brute force attacks that try to *unlock the vault*, i.e. to determine the minutiae from the vault, was analyzed based on theoretical analysis and empirical data, and reasonable parameters were deduced. However, no actual authentication system was implemented and, consequently, the False Acceptance Rate (FAR) and Genuine Accept Rate (GAR) were not determined.

Uludag et. al. [6] also used minutiae location data only, and encoded a Cyclic Redundancy Check (CRC) into the secret polynomial in order to allow verification of correctness. A drawback of their construction is that it relied on human expert for the detection of minutiae in the fingerprint image and the identification of the minutiae correspondences between fingerprints. Based on experiments, eligible parameters were determined, the FAR and GAR as well as the security against brute force attacks were determined.

In [7], Uludag and Jain refined the construction of [6] by an automatic fingerprint alignment algorithm using the locations of highest curvature of the friction ridge as additional *helper data* stored in addition to the vault. In experiments, the authors determined FAR and GAR values for a single set of parameters.

Nandakumar et. al. [8] extended the ideas of the previous constructions. Their implementation of the fuzzy fingerprint vault used both minutiae locations and orientations. Spurious or unreliable minutiae were filtered by quality indices computed from local properties of the fingerprint image, and the fingerprint alignment and minutiae matching method based on points of highest curvature of [7] was improved. Experiments were conducted on two different databases and with several sets of parameters, FAR and GAR values were reported, and the complexity of brute force attacks was estimated.

In Li et. al. [9], an alternative fingerprint alignment method for the fuzzy vault was proposed, based on the topological structures around the *core* of the fingerprint. Their implementation used both minutiae locations and orientations. Experiments were conducted, and FAR and GAR values were reported and compared to those from [8].

In [10], the authors of the present article present an implementation of the fuzzy fingerprint vault using minutiae locations of several fingers per person. Several optimizations were applied, for instance, filtering of spurious or unreliable minutiae was performed both during enrollment and during authentication by several measurements and by quality values of the feature extraction algorithm, respectively. Fingerprint alignment was performed without additional helper data but by a minutiae matcher algorithm that optimized the number of minutiae correspondences between the fingerprints by means of relative rotation and translation. A comprehensive treatment of parameter selection criteria was given with respect to security against brute force attacks, and eligible parameters deduced by combining empirical data with analytical and heuristic arguments.

Other constructions [27][28] did not use the absolute location of the minutiae at all, but features deduced from the relative topological structures around the minutiae. These features are stable with respect to orientations, and in the case of [28], even of translations. The FAR and GAR values reported in [28] are better than those from [8] at the cost of a larger template size.

In [29], Nagar et. al. propose a combination of the fuzzy fingerprint vault and the fuzzy commitment scheme. The fuzzy commitment scheme is used to individually protect the ordinate values in the vault corresponding to the minutiae, i.e. the corresponding function value of the polynomial, using *minutiae descriptors*, topological properties of the minutiae's neighborhood. Thus, an attacker has to determine both, the minutiae descriptors and their locations. The FAR and GAR values reported are much better than that of [8].

In [12], Mihailescu et. al. presented an improved brute force attack and showed that the parameters suggested by Clancy et. al. in [5], and by Uludag and Jain in [7] do not provide the claimed security.

In [30], Scheirer and Boult proposed three new attack methods beyond the scenario of reconstruction of the biometric template from a single vault. The most serious one is a correlation attack, where an attacker can retrieve the private data from combining two independently generated vaults of the same user. This attack was implemented and proved to be very efficient for relevant parameters by Kholmatov and Yanikoglu [31]. A potential countermeasure against the correlation attack was proposed by Nandakumar et. al. [32].

A complete different type of attack was proposed by Chang et. al. [11], which tried to distinguish genuine minutiae in the vault from chaff points by the number of pixels in their proximity with sufficient distance to other points in the vault. This attack seems particularly promising of the number of chaff points used is close to the maximum possible so that the minimum distance enforced between the points constitutes a dense sphere packing, as discussed in Section IV-C2.

III. The fuzzy vault for fingerprints

In this section, we define the fuzzy fingerprint vault scheme, on which our security analysis is based. This scheme basically matches the implementations presented in [5], [6], [7], [9] and [10].

A. Adaptation to fingerprints

In order to implement the fuzzy vault for fingerprints, several adaptations are necessary.

1) Selection of biometric feature: In the *fuzzy fingerprint vault*, minutiae information is used as private attributes. Minutiae are bifurcations and endings of the ridges in a fingerprint and these features are commonly used for fingerprint authentication. The error correcting capacity of the fuzzy vault scheme fits well with typical measurement errors of minutiae data, in particular with insertions, deletions, and permutations of minutiae. Whereas the constructions of the fuzzy fingerprint vault in [5], [6], [7] and [10] use minutiae locations only, [8] uses both location and orientation of the minutiae.

A somewhat surprising finding is that using minutiae orientations in addition to their locations does not add significant benefit for the privacy protection of the fuzzy fingerprint vault. According to [8] "Using minutia orientation in addition to the location attribute has two advantages. During vault encoding, it increases the number of possible chaff points that can be added because we can now add a chaff point whose location is close to a genuine template minutia but with a different direction. During vault decoding, it makes it easier to filter out the chaff points from the vault because it is less probable that a chaff point will match with the query minutia in both location and direction." While we generally agree with this statement, we stress two points. First, the number of possible chaff points, as the number of potential locations for genuine minutiae, is irrelevant for the protection of the biometric data, as an attacker only needs to determine the genuine minutiae in the vault and, thus, can neglect potential chaff points that are not stored therein. Second, the strong dependencies of a minutia orientation with the corresponding location and with the orientation of other minutiae [33][34] can facilitate distinction of genuine minutiae from chaff points in the vault. For instance in [8], an example of a vault is depicted, where many points can be visually identified as probable chaff points due to their predominantly radial directions, and many pairs of spatially close points most likely contain at least one chaff point as the orientations differ too much to be in accordance with a plausible orientation field of a fingerprint.

For these reasons we restrict our consideration to the fuzzy fingerprint vault using minutiae location information only. Nevertheless, we stress that the method proposed in [29] to utilize additional data of minutiae, e.g. their orientations, in the fuzzy vault by means of the fuzzy commitment scheme, can significantly increase security. However, analysis of this approach is beyond the scope of this paper.

2) Tolerance of minutiae mapping: In the original fuzzy vault scheme, correspondence between points in the query set and the fuzzy vault means equality. For the application of the fuzzy vault to fingerprints, the definition of minutiae correspondence is typically widened to proximity with respect to the Euclidean distance to provide tolerance with respect to small deviations in the measured minutiae locations, which are introduced by elastic skin distortions and the limitation of optical and algorithmic accuracy of the measurement. We will, therefore, assume that a minutia in a query fingerprint matches with a minutia or chaff point in the vault if both locations have an Euclidean distance of at most δ and if there is no other point in the vault closer to the minutia. In the case that several fingers are used per person, matching minutiae or chaff point must be from the same finger.

3) Fingerprint alignment: The position of a fingerprint varies between different measurements, inducing relative translations and rotations of the corresponding minutiae sets. In order to identify correspondences between the minutiae in the query fingerprints and the minutiae stored in the vault, these minutiae sets must be, at least roughly, aligned with respect to each other. The fingerprint alignment method is crucial for the fuzzy fingerprint vault, as an incorrect alignment results in a relative rotation of the query minutiae set to the points in the vault and, with very high probability, in an insufficient number of identified support points of the polynomial, which results in a failure to authenticate. Several techniques have been proposed to ensure a sufficiently correct alignment. Some implementations use topological information of the fingerprint ridge patterns [7][9], while others apply a minutiae matcher algorithm [10][8]. Early proposals [5][6] even relied on minutiae matching by human experts, which is clearly not practical. We do not impose any assumption on the method used for alignment or the goodness of the relative alignment; instead, we only consider the rate, at which the minutiae in the vault are identified in the query fingerprints (the *minutiae match rate*), which in turn depends on the alignment of these minutiae sets.

4) Combining several fingers: According to [35], the minutiae of a single finger do not provide sufficient entropy to extract a secure cryptographic key. Therefore, we allow to use minutiae from more than one finger. The minutiae of the different fingers can be easily fused on a feature level by storing with each minutia or chaff point an index of the corresponding finger. A general

discussion on approaches for multi-instance fusion in template protection schemes and implications to security can be found in [36]. Since we allow the biometric templates to be taken from several fingers of a person, the minutiae, and likewise, the chaff points, are not only represented by their location but also by the finger code. Subsequently, let both, minutiae and chaff points, be represented as points $\mathbf{m} = (\mathbf{a}, l) \in \mathbf{Z}^2 \times \{1, \ldots, f\}$, where $\mathbf{a}$ is a location in the fingerprint image represented with respect an arbitrary coordinate system, and l is an index of the finger. We will define the *distance* $||\mathbf{m}_i - \mathbf{m}_j||$ of minutiae or chaff points as the Euclidean distance of their locations $\mathbf{a}_i$ and $\mathbf{a}_j$, if both points are from the same finger, i.e., if $l_i = l_j$, and as infinite otherwise.

5) Embedding to finite field: In order to evaluate the polynomial, the minutiae data have to be embedded into the finite field. In previous implementations, the minutia location was represented as field element using a suitable encoding function. For our analysis we use a slight optimization and evaluate the polynomial not on the minutiae data but on the indices of the minutiae in the vault. This modification minimizes the size of the function values stored as part of the vault and thus the loss of entropy. See Section IV-B for details. The finite field will be chosen larger than the number of points in the vault and, consequently, injective encoding of the indices to the finite field is possible. For the ease of reading, we will omit the encoding and treat the indices as if they were field elements.

6) Storage of a hash value of the polynomial: In contrast to the implementations in [6], [7], and [8] that incorporate a CRC check sum into the polynomial's coefficients to allow verification of the recovered polynomial, we store a hash value of the coefficients, as it is done (analogously) in the fuzzy commitment scheme [19]. This approach has the advantage that it does not reduce the search space for attackers due to the internal structure of the secret. We assume that the hash value does not leak any information; this assumption is frequently used in cryptographic protocol analysis [37].

7) Selecting the feature space: Subsequently, let $\mathcal{M}$ be the set of all possible minutiae and $n = |\mathcal{M}|$ the number of possible values for a minutia or chaff point. If the fingerprint images have a resolution of $N \cdot M$ pixels (height and width), we have $\mathcal{M} \subseteq [1, N] \times [1, M] \times \{1, \ldots, f\}$, when expressing locations in Cartesian coordinates. As we will see in Section VI-A, it may be useful to restrict the set $\mathcal{M}$ to a subset with high frequency of minutiae occurrence.

B. Description of the scheme

As any biometric authentication system, the fuzzy fingerprint scheme comprises an enrollment and a verification step. In the context of the fuzzy vault, these steps are also referred to as *vault locking* and *vault unlocking*.

1) Enrollment (Locking the Vault): Let q be prime power and $k < t < r \leq q$. For each user, a random polynomial P of degree less than k over the finite field $\mathbf{F}_q$ is selected. The coefficients of this polynomial represent the secret key of the scheme. Then, a set T of t minutiae of the user is determined. This set of minutiae is amended by random chaff points, resulting in a set of r points, containing t genuine minutiae and $r - t$ chaff points. A minimum distance of d is enforced among minutiae and chaff points to reduce errors during verification by wrong mapping of close points. Furthermore, in order to ensure that minutiae and chaff points within the vault are not distinguishable by their index, they are lexicographically ordered.

For all genuine minutiae $\mathbf{m}_j$, where j is its index after applying the lexicographic order, $y_j = P(j)$ is computed. For each chaff point $\mathbf{m}_j$, where j is its index in the lexicographic order, a random value $y_j \neq P(j)$ is chosen. The *vault* consists of the lexicographically ordered list of minutiae and chaff points, paired with the corresponding y_j values. The vault and a cryptographic hash value of the concatenated coefficients of P are stored in the database.

To facilitate security analysis, we assume that the chaff points are chosen uniformly from the set $\mathcal{M}$ of potential minutiae with the restriction that the minimum distance is respected. However, since the locations of genuine minutiae are not uniformly distributed in the image area, see Section VI-A, selecting chaff points with a more natural distribution that resembles that of genuine minutiae would make them less distinguishable from the genuine minutiae in the vault. Nevertheless, since the chaff points are chosen after the genuine minutiae in the vault have been determined, those points in the vault that correspond to image locations, where minutiae occur with particularly high frequency, are more likely to be genuine minutiae anyway.

2) Verification (Unlocking the Vault): We only consider an authentication in the verification scenario, where the identity of the user is known a priori.

In order to verify the identity of a user, the minutiae are measured from a query fingerprint. Then the matches between these minutiae and the minutiae and chaff points contained in the vault are identified. Precisely, for each minutiae in the query fingerprint, the closest point in the vault with Euclidean distance smaller than a threshold δ is identified, where δ is a tolerance parameter. The matching of the minutiae in the query fingerprint with those in the vault requires a (nearly) correct alignment of the query fingerprint with respect to the minutiae in the vault. To accomplish this, either vertical alignment of the fingerprints prior to minutiae extraction , e.g., using singular point detection [38],

can be used, or a minutiae matching algorithm can be deployed that tries to find the alignment, by which the number of matches are maximized [10].

The indices of the matching minutiae and chaff points in the vault, along with the corresponding y_i values, are used to recover the secret polynomial P, see Section III-B3 for details. If the number of genuine minutiae among the matches is sufficiently high the polynomial can be recovered. See Section III-B3 for a discussion.

The correctness of the recovered polynomial is checked using the stored hash value.

3) Recovery of the polynomial: The unlocking of the vault during authentication requires the recovery of the secret polynomial from a set of points (j_i, y_{j_i}), some of which (those resulting from matches with minutiae) lying on the polynomial P, while others (those resulting from matches with chaff points) do not. For this task, an Reed-Solomon decoder is needed that on input $(j_1, y_{j_1}), \ldots, (j_\ell, y_{j_\ell}) \in \mathbf{F}_q^2$ with $\ell \geq k$, outputs $e_0, \ldots, e_{k-1} \in \{0, \ldots, q-1\}$, so that $y_{j_i} = P(j_i)$ holds for at least k of the points (j_i, y_{j_i}) with $P(z) = \sum_{i=0}^{k-1} e_i z^i$, if such a polynomial exists. We assume that the Peterson-Berlekamp-Massey-decoder is used as suggested in [3]. This technique is successful if $(\ell + k)/2$ of the x points handed over to the decoder are correct. Although there are Reed-Solomon decoders that can decode with only $\sqrt{\ell k}$ correct points, they do not offer significant advantage for the fuzzy vault, because $\sqrt{\ell k}$ is quite close to $(\ell + k)/2$ for typical parameters, and they are computationally much less efficient [3].

IV. Security analysis

It is understood that there are different threats for the fuzzy fingerprint vault and that the exposure of the original template is just one of them. Three other types of such attacks against the fuzzy vault are described in [30], among which the correlation of two vaults from independent enrollments ("record multiplicity" attack) represents a serious threat to the fuzzy vault, which is still not completely satisfactorily solved. However, a comprehensive analysis of all potential attacks against the fuzzy vault would go beyond the scope of this paper. In this contribution we focus on the security of the fuzzy fingerprint vault with respect to attacks that try to recover the template or the secret polynomial from the vault. In this context, we will investigate both lower bounds (given by information theoretical results) and upper bounds (given by known attacks) of the security.

Throughout this article, let all logarithms be to the base 2.

A. Provable Security

In this section, we provide lower bounds of security with respect to attacks that aim to recover the template or, equivalently, the secret polynomial from the vault. Precisely, these results upper bound the probability that an attacker, whatever strategy and computational resources he deploys, determines the correct polynomial or template from a given vault. The only way of the attacker to increase his success probability is to check the correctness of his output, e.g., using the hash value stored in addition to the vault, and to repeat his guessing. This "provable security" is achieved by a randomization process during enrollment, which ensures that for each given vault there are many "fitting" templates and polynomials that could have been used to generate it, and the conditional probability of any assumed template or polynomial is small.

The lower bounds on the security are given by security proofs, which are deducted from information theoretical results. We admit that the term *proof* is not completely exact here. Firstly, since the security of a biometric scheme always depends on the distribution of the biometric features within the considered population, estimations based on empirical data are necessary. Secondly, the minimum distance enforced during the enrollment constitutes a sphere packing problem that requires heuristic arguments. In the course of evaluating optimal parameters with respect to the achieved security bounds, we will use further approximations, e.g., to allow a treatment of binomial coefficients with calculus.

Following [26], we use the min-entropy $\mathbf{H}_\infty$ to quantify the security of the scheme. This measure has the advantage that it expresses the (negative logarithm of the) maximum probability of guessing, and thus, can be used to deduct lower bounds on attacks (see Theorem 1). In contrast, some publication, e.g., [39] and [40], use the Shannon entropy $\mathbf{H}$ to assess the security of biometric template protection. The use of the Shannon entropy might be appealing due to the rich underlying mathematical theory, which allows to deduct quite impressive results, e.g., see [41]. However, as shown in [42], the Shannon entropy can, for certain probability distributions, be very insignificant for assessing the minimum attack complexities. In general, the inequality $\mathbf{H}_\infty(A) \leq \mathbf{H}(A)$ holds for any random variable A, but, in the opposite direction, the Shannon entropy can exceed the min-entropy (and thus the logarithm of attack complexities) by any factor. Consequently, the Shannon entropy is not the eligible measure to determine the capacity of the fuzzy fingerprint vault with respect to provable security, i.e., to lower bounds for attack complexities.

Subsequently, let P(X) denote the probability of an event X and let $\mathbf{E}_{a \leftarrow A}[f(a)]$ be the expectation of the function value of a random variable A. The *min-entropy*

of a random variable A is given by

$$\mathbf{H}_\infty(A) := -\log(\max_a(\mathrm{P}(A=a))),$$

and the *average min-entropy* of A given B is defined as

$$\widetilde{\mathbf{H}}_\infty(A \mid B) := -\log\left(\mathbf{E}_{b\leftarrow B}\left[2^{-\mathbf{H}_\infty(A \mid B=b)}\right]\right)$$

For a biometric encryption scheme with feature vector T and vault Y, we call $\mathbf{H}_\infty(T) - \widetilde{\mathbf{H}}_\infty(T \mid Y)$ the *loss of entropy.*

B. Minimum attack complexity

The following result shows that the security of the fuzzy vault for fingerprints can be lower bounded by the average min-entropy of the biometric feature vector given the vault. The result is a trivial adaptation of Theorem 1 and Lemma 2 from [23], and follows immediately from the definition of the min-entropy. It holds (with according notations) for any biometric encryption scheme, in which the secret key and the vault uniquely determine the biometric feature vector.

Theorem 1. *Any algorithm that takes as input the vault Y and tries to output the secret polynomial $P(x) = \sum_i e_i x^i$ or the set of minutiae T has an average success probability of at most $2^{-\widetilde{\mathbf{H}}_\infty(T \mid Y)}$.*

An attacker who has determined the original template T of a user can recover the secret polynomial P by simulating a verification using T and the vault Y; the stored hash value allows checking, if the resulting polynomial is correct. On the other hand, if an attacker has (somehow) learned P, he can easily recover T from the vault Y, simply by determining all $\mathbf{m}_j$ in Y with $y_j = P(j)$. Therefore, it is equally difficult to recover the template T as to determine the secret polynomial P. In terms of information theory, we obtain the following result:

Theorem 2. *Given the stored reference data (vault and hash value), recovering to biometric template T is computationally equivalent to determination the secret polynomial P. Moreover, $\widetilde{\mathbf{H}}_\infty(T \mid Y) = \widetilde{\mathbf{H}}_\infty(P \mid Y)$.*

On the other hand, the success probability of an *fingerprint dictionary attack* (see Section IV-E) trying to recover the polynomial by choosing random templates equals, by definition, the False Accept Rate (FAR), while the min-entropy upper bounds the probability of any attack. Therefore, we can state the following result, which was presented already - for a larger class of schemes and using different mathematical notations - in [43].

Theorem 3. $\widetilde{\mathbf{H}}_\infty(P \mid Y) \leq -\log(\mathrm{FAR})$.

Theorem 3 implies that the information content of a cryptographic key extracted from P cannot exceed $-\log(\mathrm{FAR})$. In [35], this conclusion was drawn for arbitrary schemes, in which biometric data is used to extract a secret key. Since it is indeed possible to recover P from Y in 1/FAR steps on average by the fingerprint dictionary attack (see Section IV-E), the length of any cryptographic key secured by a fuzzy fingerprint vault should not exceed $-\log(\mathrm{FAR})$ bits. In order to extract this number of bits from P while preserving all its entropy, it can be used as a seed of a pseudo-random number generator.

C. Loss of entropy

By definition, the average min-entropy of the biometric feature vector given the vault is the difference between the entropy of the feature vector and the loss of entropy. We now turn to the estimation of the latter quantity. We first consider the case, where no minimum distance is enforced among the minutiae and chaff points, i.e., the case $d = 1$, and then generalize these results to the case $d > 1$.

1) The case of trivial minimum distance: In [13], Lemma D.1, a lower bound for the loss of entropy in the original fuzzy vault scheme has been given. In the case $d = 1$, i.e., if the minimum distance is trivial and the minutiae and chaff points only need to be distinct, the result can be directly applied to our implementation. The proof is a simple adaptation of the proof of Lemma D.1 in [13].

Theorem 4. *If $d = 1$, the loss of entropy is at most $(t-k)\log q - \log\binom{r}{t} + \log\binom{n}{t} + 2$, i.e.,*

$$\widetilde{\mathbf{H}}_\infty(T \mid Y) \geq \mathbf{H}_\infty(T) - (t-k)\log q + \log\binom{r}{t} - \log\binom{n}{t} - 2. \quad (1)$$

Proof: By Lemma 3.1 in [13]

$$\widetilde{\mathbf{H}}_\infty(T \mid Y) \geq \mathbf{H}_\infty(T, Y) - \lambda,$$

where 2^λ is the number of possible values that Y can take.

We first estimate $\mathbf{H}_\infty(T, Y)$. The information contained in T and Y is composed of four parts: The set of minutiae T, the set of chaff points, the y_i-values for the minutiae, and the y_i-values for the chaff points. The entropy of the $r - t$ chaff points is given by $\log\binom{n-t}{r-t}$, because they are randomly selected from all $n-t$ potential points that are distinct from the t minutiae. Given T, there is a one-to-one correspondence between the y_i-values for the minutiae and the random polynomial P; hence, their entropy is $k\log q$. Finally, the y_i-values for the chaff points are randomly selected from $\mathbf{F}_q \setminus \{P(i)\}$, and therefore their entropy is $(r-t)\log q - 1$. This sums up to

$$\mathbf{H}_\infty(T, Y) = \mathbf{H}_\infty(T) + \log\binom{n-t}{r-t} + k\log q + (r-t)\log(q-1).$$

On the other hand, since the minutiae and chaff points in Y are in lexicographic order, we have $2^\lambda = \binom{n}{r} q^r$. Using $(r-t)\log\frac{q}{q-1} < q\log\frac{q}{q-1} \leq 2$ and

$$\binom{n}{r}\binom{r}{t} = \binom{n}{t}\binom{n-t}{r-t}$$

this yields the result. Q.E.D.

This result can be interpreted as follows:

- The term $(t-k)\log q$ represents the information leaked by the redundantly encoded secret polynomial. Precisely, this term is composed of the $t\log(q)$ bits of information revealed by the y_i-values corresponding to the genuine minutiae and the $k\log(q)$ of information contained in the secret polynomial.
- The term $\log\binom{r}{t}$ estimates the amount of security contributed by "hiding" the t genuine minutiae among the r chaff points.
- The term $\log\binom{n}{t}$ refers to the information leaked by publishing T as part of the vault.

Since $\mathbf{H}_\infty(T) \leq \log\binom{n}{t}$, the lower bound (1) is positive (and hence meaningful) only if $q^{t-k} \leq \binom{r}{t} \leq \binom{q}{t} \leq q^t/(t!)$, which implies $q > (t/\mathrm{e})^{t/k}$. The exponent t/k defines the error correction capacity of the scheme and, according to our experiments, must be larger than 1.5 to achieve a satisfactory false rejection rate (FRR). Therefore, we can obtain a scheme with provable security according to Theorems 1 and 4 only if q is considerably greater than $(t/\mathrm{e})^{1.5}$.

The bound provided by Theorem 4 is not tight. In particular, in the estimation of λ, the number of possible values for $(y_1,\ldots,y_r)$ is smaller than q^r, because, by construction, $(y_1,\ldots,y_r)$ can only assume those vectors, for which at least t of the pairs (i,y_i) lie on a common polynomial of degree smaller than k. This is exactly the set of words in the Reed-Solomon code $\mathrm{RS}_q(r,k)$ having error distance, i.e., Hamming distance to the next codeword, at most $r-t$. We are not aware of any estimation on their number that could be used to improve Theorem 4. On the other hand, for $t < (r+k)/2$ the Hamming spheres of radius $r-t$ around the code words overlap and hence already cover a significant part of $\mathbf{F}_q^r$. Thus, for $t \ll (r+k)/2$ it is not clear whether a better estimate for λ would result in a significant improvement of Theorem 4.

If we chose the chaff point according to a distribution that resembles that of minutiae locations (instead of uniformly from $\mathcal{M} \setminus T$), we would end up with a smaller bound for $\widetilde{\mathbf{H}}_\infty(T\,|\,Y)$. This reduction of provable security is paradox, as a more natural distribution of the chaff points makes them less distinguishable from the genuine minutiae, and hence, strengthens the security. However, the proof techniques used in Lemma D.1 of [13] measure the information leakage not by the entropy of Y but by the number of its possible values. Therefore, a non-uniform distribution does not change the estimate of the leaked information while it reduces the entropy added.

2) The case of non-trivial minimum distance: For the case $d \geq 2$, we have to analyze the effect of the minimum distance to the number of possible choices for the chaff points and the possible values for the vault Y. Subsequently, we will use the following definitions.

For a point $\mathbf{m} \in \mathcal{M}$ let $B_d(\mathbf{m})$ denote the set of points in $\mathcal{M}$ that have Euclidean distance to $\mathbf{m}$ smaller than d, and let $V_d = 1 + 4\sum_{i=1}^{\lceil d-1\rceil}\lceil\sqrt{d^2-i^2}\rceil$ be the number of integer points $\mathbf{m} \in \mathbf{Z}^2$ with Euclidean norm smaller than d. Obviously, $|B_d(\mathbf{m})| \leq V_d$.

Since the minutiae and chaff points are selected with minimum distance d, the d-sphere centered at a selected point is excluded from the potential values for subsequent points. If the d-sphere neither juts out beyond $\mathcal{M}$ nor intersects with the d-spheres of the previously selected points, the number of possible choices for the next point is reduced by exactly V_d; otherwise, the reduction is smaller.

These effects make an exact estimation of the number of possible choices for the chaff points or the number of potential values for Y virtually impossible. However, for $rV_d \ll n$, the likelihood that a selected point is too close to the boundary of $\mathcal{M}$ or to a previously selected point is small. In this case, the approximation that, on average, each point reduces the number of choices for the subsequent points by V_d is quite accurate. Subsequently, we assume $rV_d \ll n$, and thus, approximate the number of chaff points by $V_d^{r-t}\binom{n/V_d-t}{r-t}$ and the number of possible values for Y by $V_d^r\binom{n/V_d}{r}$. Analogously to Theorem 4, we obtain the following result:

Theorem 5. *For $rV_d \ll n$, the maximal loss of entropy is approximately $(t-k)\log q - \log\binom{r}{t} + \log\binom{n/V_d}{t} + t\log V_d + 2$, i.e., $\widetilde{\mathbf{H}}_\infty(T\,|\,Y) \geq E$ with*

$$E \approx \mathbf{H}_\infty(T) - (t-k)\log q + \log\binom{r}{t} - \log\binom{n/V_d}{t} - t\log V_d - 2. \quad (2)$$

D. Entropy of the feature vector

The entropy of the feature vector T is defined by the maximum likelihood that it takes a certain instance M. Since for the parameters of interest the number of possible instances by far exceeds the number of persons, for which minutiae information is available, we can estimate the entropy of T only by modeling its probability distribution. Several publications have proposed models for minutiae distributions, e.g., [44] and [45]. However, their analysis already takes into consideration the error

tolerance of the minutiae matching algorithm and is therefore not applicable for the determination of the raw entropy $\mathbf{H}_\infty(T)$.

We model the probability distribution of T by a probabilistic process **Select_T**, where the t minutiae are successively chosen. The first minutia $\mathbf{m}_1$ is selected according to a distribution $\mathcal{D}$ defined over $\mathcal{M}$. All subsequent minutiae $\mathbf{m}_i$ are selected to the same distribution $\mathcal{D}$ restricted to the areas in $\mathcal{M}$ not covered by the d-spheres $B_d(\mathbf{m}_1), \ldots, B_d(\mathbf{m}_{i-1})$ around the previously chosen minutiae.

Like all previous models for the distribution of minutiae, we do not assume any statistical dependency between the locations of the individual minutiae, except that they have the minimum distance d. Although it is known that minutiae tend to overdisperse on a small scale (precisely, between 11 and 20 pixels for 500 dpi) and to cluster on a large scale [46]. The overdispersion on a small scale can be partially explained by minimum distances typically enforced by minutiae extraction algorithms to avoid ambiguous results, e.g., see [47], but in [46] biological arguments taken from [48] are used. Due to the enforcement of a minimum distance d during template selection this effect is in line with our model, at least for sufficiently large d. Furthermore, the overdispersion reported in [46] is rather weak. On the other hand, in [46] the observed clustering on a large scale is explained by a higher minutiae frequency around core or delta points. This effect is addressed in our model by using a non-uniform distribution $\mathcal{D}$, in which higher probabilities refer to such cluster points. Of course, there could be more complex dependencies between the location of individual minutiae. However, to our knowledge, there are no observations or models implying such dependencies (we refer to [49] for a detailed discussion of this aspect).

Using our statistical model, we can show the following result:

Theorem 6. *If T is chosen according to the random process* **Select_T** *and the maximum likelihood of a minutiae location is $1/\psi$, then*

$$\mathbf{H}_\infty(T) \geq \log \binom{\psi/V_d}{t} + t \log V_d$$

Proof: Let $\mathrm{P}(A)$ denote the probability of random event A. Furthermore, for $i = 1, \ldots, t$ let M_i let be the random variable of the i-th point output by **Select_T**. By M we denote the random variable chosen according to $\mathcal{D}$. Then by definition

$$\begin{aligned} 2^{-\mathbf{H}_\infty(T)} &= \max\left(\mathrm{P}\left(\{M_1, \ldots, M_t\} = \{\mathbf{m}_1, \ldots, \mathbf{m}_t\}\right)\right) \\ &\leq t! \max\left(\mathrm{P}\left(M_1 = \mathbf{m}_1, \ldots, M_t = \mathbf{m}_t\right)\right), \quad (3) \end{aligned}$$

where the maximum is taken over all $\mathbf{m}_1, \ldots, \mathbf{m}_t$. The latter probability $\mathrm{P}\left(M_1 = \mathbf{m}_1, \ldots, M_t = \mathbf{m}_t\right)$ can be expanded to

$$\prod_{i=1}^{t} \mathrm{P}\left(M_i = \mathbf{m}_i \,|\, \forall j < i : \; M_j = \mathbf{m}_j\right).$$

The first term has an empty condition and is limited by $1/\psi$, while the other factors can be estimated as follows:

$$\begin{aligned} &\mathrm{P}\left(M_i = \mathbf{m}_i \,|\, M_1 = \mathbf{m}_1, \ldots, M_{i-1} = \mathbf{m}_{i-1}\right) \\ &\quad = \mathrm{P}\left(M = \mathbf{m}_i \,|\, M \notin B_d(\mathbf{m}_1) \cup \ldots \cup B_d(\mathbf{m}_{i-1})\right) \\ &\quad = \frac{\mathrm{P}\left(M = \mathbf{m}_i \wedge M \notin B_d(\mathbf{m}_1) \cup \ldots \cup B_d(\mathbf{m}_{i-1})\right)}{\mathrm{P}\left(M \notin B_d(\mathbf{m}_1) \cup \ldots \cup B_d(\mathbf{m}_{i-1})\right)} \\ &\quad \leq \frac{\mathrm{P}\left(M = \mathbf{m}_i\right)}{1 - \mathrm{P}\left(M \in B_d(\mathbf{m}_1) \cup \ldots \cup B_d(\mathbf{m}_{i-1})\right)} \end{aligned}$$

By assumption, the numerator is at most $1/\psi$, while the probability in the denominator is limited by the term $\left|B_d(\mathbf{m}_1) \cup \ldots \cup B_d(\mathbf{m}_{i-1})\right| / \psi$, which is at most $(i-1)V_d/\psi$. This results in

$$\begin{aligned} &\mathrm{P}\left(M_i = \mathbf{m}_i \,|\, M_1 = \mathbf{m}_1, \ldots, M_{i-1} = \mathbf{m}_{i-1}\right) \\ &\qquad \leq \frac{1}{\psi - i \cdot V_d} \quad (4) \end{aligned}$$

Consequently, with (3) we obtain

$$2^{-\mathbf{H}_\infty(T)} = t! \prod_{i=0}^{t-1} \frac{1}{\psi - i \cdot V_d}.$$

The desired result now follows by elementary transformations. Q.E.D.

By combining Theorem 5 with Theorem 6 we obtain the following Theorem.

Theorem 7. *For $d \geq 1$, $\widetilde{\mathbf{H}}_\infty(T \,|\, Y) \geq E$ with*

$$E \approx \log \binom{\psi/V_d}{t} - (t-k) \log q + \log \binom{r}{t} - \log \binom{n/V_d}{t} - 2,$$

where $1/\psi$ is the maximum likelihood of a minutiae location.

E. Practical Security

In the previous sections, we have focused on provable security in terms of lower bounds for the number of trials for attacks. However, these bounds are not at all sharp, as existing attacks are much less efficient than these bounds would allow. For this reason, we now consider practical attacks and analyze the security of the fuzzy vault with respect to these attacks.

For the recovery of the original template and the secret polynomial from a single vault two kinds of brute force attacks can be distinguished: exhaustive search on the templates or exhaustive search on the polynomials.

1) Fingerprint dictionary attack: In the *fingerprint dictionary attack* an attacker collects a large number of realistic templates, either from real fingerprints or artificially. For all these templates he simulates the authentication procedure using the vault until the secret polynomial has been found. Although it has been shown in [3] that for typical parameters there is with high probability a large number of polynomials "fitting" the vault, i.e., there are many polynomials of degree smaller than k such that exactly t of the stored points lie on the polynomial, the attacker can check the correctness of the polynomial using the stored hash value (see Section III-B2). If the templates are chosen with the same probability distribution as they occur within the group of users of the biometric application, the success probability of each attempt equals the False Accept Rate (FAR) and the attacker needs FAR^{-1} trials on average. (For this reason, this attack is also referred to as *FAR attack* [36].) This results in an overall workload of $N_\mathrm{v} \cdot \mathrm{FAR}^{-1}$, where N_v is the effort for a single verification.

Usually, the FAR is determined empirically by performing a sufficiently large number of impostor matches, i.e., matches with fingerprints of other individuals. Of course, the empirical determination has to be done for every set of parameters separately. Subsequently, we explore if the FAR can also be estimated theoretically as a function in dependence of the parameters of the scheme.

Let m_c denote the number of *correct matches*, i.e., the matches between the query fingerprint and the genuine minutiae, and let m_f be the number of *false matches*, i.e., the matches between the query fingerprint and the chaff points. According to Section III-B3, the reconstruction of the polynomial is only possible if $m_\mathrm{c} \geq m_\mathrm{f} + k$. Therefore, we obtain

$$\mathrm{FAR} = \sum_{a=k}^{t} \sum_{m_\mathrm{f}=0}^{b-k} \mathrm{P}(m_\mathrm{c} = a \wedge m_\mathrm{f} = b).$$

The probability $\mathrm{P}(m_\mathrm{c} = a \wedge m_\mathrm{f} = b)$ that there are exactly a correct and b false matches depends on the specific method used to identify minutiae correspondences between the query fingerprint and the vault. This method usually searches for the correct relative alignment of the minutiae set in order to compensate global rotations and translations of the fingerprints. The precision and reliability of the alignment method has great impact on the probability $\mathrm{P}(m_\mathrm{c} = a \wedge m_\mathrm{f} = b)$. Therefore, a reasonably accurate theoretical estimation of the FAR is unfeasible unless very simplifying assumptions are made on the alignment, e.g., that the fingerprints are perfectly aligned.

2) Polynomial reconstruction: An attacker can try to recover the polynomial from the stored points directly, i.e., without exploiting knowledge about the distribution of minutiae and the corresponding feature vectors. The underlying computational problem is known as Reed-Solomon decoding problem or *polynomial reconstruction* problem. It is believed to be hard for $k < t < \sqrt{r(k-1)}$, and it is known that random instances of this problem are as hard as the worst case [50]. For very large fields sizes, it is known to be NP-complete [51]. For these reasons, it has been repeatedly suggested as a basis for cryptographic constructions [50].

According to [52], two approaches are most efficient for the polynomial reconstruction: Either, after guessing k genuine minutiae, the polynomial is reconstructed using polynomial interpolation, e.g., by Lagrange interpolation, or it is determined by Reed-Solomon list decoding after guessing $\Delta = r - \frac{t^2}{k-1} + 1$ of the chaff points. Let aside the fact that polynomial interpolation is much more efficient than Reed-Solomon list decoding, for typical parameters (and all parameters suggested so far), $\Delta > k$ and therefore, it is more efficient to guess k genuine minutiae among the stored points.[1] This approach has been used by the attack of Mihailescu, et. al. [12], which systematically searches through all subsets $\{j_1, \ldots, j_k\}$ of $\{1, \ldots, r\}$, computes the unique polynomial P satisfying $P(j_i) = y_{j_i}$ by polynomial interpolation, and checks the correctness of this polynomial. Assuming (as done in [12]) that all points in the vault are equally likely to be a genuine minutia, this attack needs $\binom{r}{k}/\binom{t}{k}$ trials on average. In [12], the number of operations needed for each interpolation is estimated as $6.5k\log^2(k)$ using results from [53]. However, this estimation is incorrect, as 6.5 is the explicit constant for the running time of fast polynomial interpolation only if it is expressed in terms of the running time $\mathsf{M}(k)$ for multiplication of polynomials of degree k (see [53], Corollary 10.2). Dissolving $\mathsf{M}(k)$ to $O(k\log(k))$ introduces another factor of 18 (see Corollary 8.19 in [53]). Thus, we have to correct the running time estimation for polynomial interpolation used by [12] to $117k\log^2(k)$. This results in an average number of

$$W \leq 117k\log^2(k)\binom{r}{k}\binom{t}{k}^{-1}, \tag{5}$$

operations required for the attack. Note that the term "operations" refers to additions, subtractions, multiplication and division over $\mathbf{F}_q$.

Of course, the assumption that all points in the vault are genuine minutiae with the same probability, is an

[1] This is in contrast to the (obviously wrong) statement in [52].

oversimplification. There are (at least) two effects resulting in a non-uniform distribution of these probabilities, which are subsequently discussed.

Firstly, it has been shown in [11] that since the chaff points are selected after the genuine minutiae in the vault were determined, the average free area (not occupied by the d-spheres $B_d(\mathbf{m})$ of other points) in the proximity of chaff points is smaller than that around genuine minutia. This tendency can be exploited to tell apart genuine minutiae from chaff points more efficiently than by mere guessing. In [11], the method has been shown to be efficient in the case of a maximum number of chaff points; given a density 0.45 for random sphere packings [5], the maximum number r of points in the vault is $0.45 \cdot n/V_{\lceil d/2 \rceil}$. In this case, the polynomial reconstruction attack can be sped up considerably by preferring those points having more free area in their neighborhood than others. However, it has been shown in [5] that if the number of chaff points is considerably smaller than their maximum, the effect exploited by the attacker is much weaker. We assume that r is chosen considerably smaller than its maximum value, i.e., that $r \ll 0.45 \cdot n/V_{\lceil d/2 \rceil}$, and thus, this attack method is less efficient than the second approach for selecting points in the polynomial reconstruction (see following paragraph). We will critically review this assumption on the basis of our results in Section VI-D.

Secondly, the locations of minutiae are not uniformly distributed. Even if chaff points were selected using a "natural" distribution, i.e., a distribution resembling that of minutiae, they were less likely to occupy frequent minutiae locations than the genuine minutiae in the vault, because the latter ones are selected before the chaff points are chosen. As we assumed that chaff points are selected according to a uniform distribution, this effect is even stronger. Subsequently, we discuss the advantage an attacker can gain from this effect.

Obviously, the best strategy for speeding up the polynomial reconstruction is to try points $\mathbf{m}_i \in R$ with higher conditional probability $\mathrm{P}(\mathbf{m}_i \in T \,|\, \mathbf{m}_i \in R)$ first. In particular, an optimized attack would first determine the minutiae occurrence frequency for all locations in $\mathcal{M}$, sort the points $\mathbf{m}_1, \ldots, \mathbf{m}_r \in R$ according to the frequency p_i corresponding to their location so that $p_1 \geq \ldots \geq p_r$, and would then search through all subsets $\{j_1, \ldots, j_k\}$ of $\{1, \ldots, K\}$ with increasing $K \geq k$. We call this optimized attack the *smart polynomial reconstruction*. Up to our knowledge, this (quite obvious) improvement of the polynomial reconstruction attack on the fuzzy fingerprint vault has not been proposed in the literature so far, although in [32], the basic idea "to exploit the non-uniform nature of biometric features and develop attacks based on statistical analysis of points in the vault" has already been phrased.

The following results enables us to deduct an approximate upper bound for the success probability of this attack method. As for the proof of Theorem 5, we assume $rV_d \ll n$ and use the approximation that, on average, each point selected for R reduces the number of choices for the subsequent points by V_d. We will critically review this assumption on the basis of our results in Section VI-D.

Lemma 8. *Let $rV_d \ll n$ and $\mathbf{m}_{j_1}, \ldots, \mathbf{m}_{j_k}$ be an arbitrary subset of points from the vault R. Then, the probability p that these points are all genuine minutiae, is approximately upper bounded by*

$$p \lessapprox \prod_{i=1}^{k} \frac{(t-i+1)(n-tV_d)}{(r-t)\psi_i + (t-i+1)n - t(r-i+1)V_d},$$

where $1/\psi_i$ is the probability of a minutiae occurrence at position $\mathbf{m}_{j_i}$.

Proof: For the ease of reading, we use $\mathrm{P}^{(i-1)}(A)$ to denote a probability of an event A under the condition that points $\mathbf{m}_{j_1}, \ldots, \mathbf{m}_{j_{i-1}}$ are genuine minutiae, i.e.,

$$\mathrm{P}^{(i-1)}(A) = \mathrm{P}(A \,|\, \mathbf{m}_{j_1}, \ldots, \mathbf{m}_{j_{i-1}} \in T).$$

The probability p that the points $\mathbf{m}_{j_1}, \ldots, \mathbf{m}_{j_k}$ are all genuine minutiae is given by

$$p = \prod_{i=1}^{k} \mathrm{P}^{(i-1)}(\mathbf{m}_{j_i} \in T \,|\, \mathbf{m}_{j_i} \in R). \tag{6}$$

We can estimate

$$\begin{aligned} &\mathrm{P}^{(i-1)}(\mathbf{m}_{j_i} \in T \,|\, \mathbf{m}_{j_i} \in R) \\ &\qquad = \frac{\mathrm{P}^{(i-1)}(\mathbf{m}_{j_i} \in T)}{\mathrm{P}^{(i-1)}(\mathbf{m}_{j_i} \in R)} \\ &\qquad = \left(1 + \frac{\mathrm{P}^{(i-1)}(\mathbf{m}_{j_i} \in R \setminus T)}{\mathrm{P}^{(i-1)}(\mathbf{m}_{j_i} \in T)}\right)^{-1} \end{aligned} \tag{7}$$

For the estimation of $\mathrm{P}^{(i-1)}(\mathbf{m}_{j_i} \in T)$, we again assume that the t genuine minutiae in T are successively selected by the probabilistic process **Select_T** introduced in Section IV-D.

Since $rV_d \ll n$, we can approximate that, on average, each point selected for R reduces the number of choices for the subsequent points by V_d. This estimation is actually an upper bound, which holds independently from our assumption $rV_d \ll n$, and we will use this bound to obtain an exact upper bound for the denominator in (7). For the nominator, however, we need a lower bound, and therefore, we use this estimation on the reduction of potential points as an approximation and not as a bound.

By (4), the probability that minutia $\mathbf{m}_{j_i}$ is chosen as the i-th minutia if $i-1$ genuine minutiae are already fixed is at most $(\psi_i-(i-1)V_d)^{-1}$. If $\mathbf{m}_{j_i}$ is not chosen as the i-th minutia, it can be selected as $(i+1)$-th minutiae only, if the i-minutiae has not been chosen from $B_d(\mathbf{m}_{j_i})$. Thus, the probability, that $\mathbf{m}_{j_i}$ is selected as $(i+1)$-th minutia, given that the first $i-1$ minutiae in T are fixed, is at most

$$\left(1-\frac{V_d}{\psi_i-(i-1)V_d}\right)\frac{1}{\psi_i-iV_d}=\frac{1}{\psi_i-(i-1)V_d}.$$

Analogously, for all $m \geq i$, the probability that $\mathbf{m}_{j_i}$ is selected as m-th minutia in T given that the first $i-1$ minutiae are fixed can be upper bounded by $(\psi_i-(i-1)V_d)^{-1}$. Thus, we obtain

$$\mathrm{P}^{(i-1)}(\mathbf{m}_{j_i} \in T) \leq \frac{t-i+1}{\psi_i-(i-1)V_d}. \quad (8)$$

For the estimation of the denominator of (7), we observe that $\mathbf{m}_{j_i}$ can be a chaff point only if none of the points in its d-sphere $B_d(\mathbf{m}_{j_i})$ is in T, i.e., if $B_d(\mathbf{m}_{j_i}) \cap T = \emptyset$. Thus, we have

$$\mathrm{P}^{(i-1)}(\mathbf{m}_{j_i} \in R \setminus T) = \mathrm{P}^{(i-1)}(Z) \cdot \mathrm{P}(\mathbf{m}_{j_i} \in R \,|\, Z), \quad (9)$$

where we have abbreviated the event $B_d(\mathbf{m}_{j_i}) \cap T = \emptyset$ by Z. Note, that the second probability does not depend on the condition that points $\mathbf{m}_{j_1}, \ldots, \mathbf{m}_{j_m}$ are genuine minutiae, because the specific configuration of T is irrelevant for the chances of $\mathbf{m}_{j_i}$ being selected as chaff point, as long as these points do not lie within the d-sphere of $\mathbf{m}_{j_i}$.

Using (8), we obtain

$$\begin{aligned}\mathrm{P}^{(i-1)}(Z) &\geq 1-\frac{(t-i+1)V_d}{\psi_i-(i-1)V_d}\\ &= \frac{\psi_i-tV_d}{\psi_i-(i-1)V_d}. \quad (10)\end{aligned}$$

On the other hand, we have

$$\mathrm{P}\left(\mathbf{m}_{j_i} \in R \,|\, Z\right) = \sum_{l=1}^{r-t} \mathrm{P}(M_l = \mathbf{m}_{j_i} \,|\, Z), \quad (11)$$

where M_l is the random variable defined by the selection of the l-th chaff point. Since we assume that the chaff points are chosen according to a uniform distribution and that each point selected for R reduces the number of choices for the subsequent points by V_d, the probability that the first chaff point chosen is $\mathbf{m}_{j_i}$, provided that none of the points in its d-sphere is in T, is approximately $(n-tV_d)^{-1}$.

If $\mathbf{m}_{j_i}$ is not chosen as the first chaff point, it can be selected as second chaff point only, if the first chaff point has not been chosen from $B_d(\mathbf{m}_{j_i})$. Thus, the probability, that $\mathbf{m}_{j_i}$ is selected as second chaff point, given that none of the points in its d-sphere is in T, is approximately

$$\left(1-\frac{V_d}{n-tV_d}\right)\frac{1}{n-(t+1)V_d}=\frac{1}{n-tV_d}.$$

Analogously, for all $m \geq 1$, the probability that $\mathbf{m}_{j_i}$ is selected as m-th chaff point, given that none of the points in its d-sphere is in T, can be approximated by $(n-tV_d)^{-1}$. Thus, we obtain from (11)

$$\mathrm{P}(\mathbf{m}_{j_i} \in R \,|\, Z) \approx \frac{r-t}{n-tV_d}, \quad (12)$$

and with (10) and (9)

$$\mathrm{P}^{(i-1)}(\mathbf{m}_{j_i} \in R \setminus T) \gtrapprox \frac{\psi_i-tV_d}{\psi_i-(i-1)V_d} \cdot \frac{r-t}{n-tV_d}. \quad (13)$$

Combining (7) with (8) and (13), we get

$$\begin{aligned}&\mathrm{P}^{(i-1)}(\mathbf{m}_{j_i} \in T \,|\, \mathbf{m}_{j_i} \in R)\\ &\qquad \lessapprox \frac{(t-i+1)(n-tV_d)}{(r-t)\psi_i+(t-i+1)n-t(r-i+1)V_d},\end{aligned}$$

and with (6) this yields the desired result. Q.E.D.

As in Theorem 6, let $1/\psi$ be the maximum likelihood of a minutiae location within $\mathcal{M}$. Then, by Lemma 8, the success probability for each individual trial of the smart polynomial reconstruction is approximately limited by

$$\prod_{i=1}^{k} \frac{(t-i+1)(n-tV_d)}{(r-t)\psi+(t-i+1)n-t(r-i+1)V_d}.$$

This general bound, together with the estimation of $117k\log^2(k)$ operations for a polynomial interpolation, provides an approximate lower bound for the average number of operations needed for the smart polynomial reconstruction.

Theorem 9. *For $rV_d \ll n$, the expected number W of operations for the smart polynomial reconstruction is $W = 117k\log^2(k)S$, where S is approximately lower bounded by*

$$S \gtrapprox \prod_{i=1}^{k} \frac{(r-t)\psi+(t-i+1)n-t(r-i+1)V_d}{(t-i+1)(n-tV_d)} \quad (14)$$

Note, that the estimation (14) depends on the number f of fingers used per person, as both n and ψ scale linearly with f. Precisely, the estimate for S decreases with f, i.e., the estimated workload of the attack is minimal for $f=1$. This dependency may surprise at first sight, but indeed, with increasing number of minutiae per finger they become more distinguishable from chaff points because less space in the area frequently assumed by minutiae is left for chaff points.

Estimating the probability ψ_i of minutiae occurrence at location $\mathbf{m}_{j_i}$ by ψ is of course quite rough. With

increasing K, the average of ψ_i for $j_i \leq K$ will decrease and so does the success probability. However, the rate of this decrease depends on the specific distribution of the minutiae locations and can only be determined on the basis of extensive data evaluation, which would go beyond the scope of this paper. Therefore, in Section VI-D, we will use Theorem 9 as a lower bound for workload of the smart polynomial reconstruction attack and complement this estimation by using the expected run time (5) of the conventional polynomial reconstruction as an upper bound.

3) Discussion: As explained in Section IV-E1, we are not able to provide a reasonably accurate run time estimation for the fingerprint dictionary attack, because theoretical analysis of the FAR is not possible without very simplifying assumptions. Therefore, the FAR needs to be determined empirically for each set of parameters used. Unfortunately, determination of very small FAR values is computationally very expensive: while the FAR for the multi-finger setting can be extrapolated from the FAR of a single-finger setting, determination of latter one requires considerably more than FAR^{-1} matching operations. Since security of the fuzzy fingerprint vault against the fingerprint dictionary attack requires a very low FAR, this task can be quite challenging.

A potential advantage of the fingerprint dictionary attack over polynomial reconstruction is that it takes optimal advantage of the actual statistical distribution of the feature vectors in the considered population. While the smart polynomial reconstruction attack exploits the non-uniformity of the minutiae locations in the considered area $\mathcal{M}$, the fingerprint dictionary attack can also take advantage from statistical dependencies among the minutiae locations. However, as discussed in Section IV-D, the dependencies reported in the literature are quite weak.

On the other hand, the effort N_{v} for each trial in the fingerprint dictionary attack is computationally expensive, as it comprises feature extraction, minutiae set alignment and Reed-Solomon decoding. For instance, we have implemented a matching algorithm that aligns the set of minutiae from the query fingerprints with the vault by determining the rotation and translation for optimal alignment (see [10] for details). For typical parameters the matching using this algorithm needs between 0.3 and 1 second on a standard PC. Of course, the matching process could be accelerated by using more sophisticated methods, but the alignment is definitely a complex task, which consumes considerable time. Moreover, the extraction of minutiae from (real or artificial) fingerprints requires extensive image pre-processing and edge detection, which is also very time consuming. In contrast, the run time estimations (5) and (14) of the polynomial reconstruction attacks counts elementary operations, i.e., additions, subtractions, multiplication or division over $\mathbf{F}_q$. In an implementation of the polynomial reconstruction attack reported in [12], the number of polynomials interpolated and tested per second of CPU time on a standard PC was greater than 8000 for $k = 14$. Based on the estimate $117k\log^2(k)$ for the number of operations needed per polynomial interpolation, and an optimistic estimate of 0.25 seconds of CPU time for a feature extraction and matching operation, we can roughly estimate that a single trial in the fingerprint dictionary attack takes 50 million times more computation time than the finite field operations counted in (5) and (14).

In this paper, we will subsequently estimate the practical security of the fuzzy fingerprint vault by the workload W of smart polynomial reconstruction attack, for which we use (5) as an upper and (14) as a lower bound. Nevertheless, we stress, that a security assessment of a concrete implementation of the fuzzy fingerprint vault should also comprise an empirical evaluation of the FAR and a resulting estimation for the workload of the fingerprint dictionary attack.

V. Optimization of Parameters

In this section we try to determine criteria for the optimal selection of parameters for both provable security and security against existing attacks. Furthermore, we derive estimates on the achievable security according to Theorems 1 and 7. We do this by estimating the maximum of E over t, k and r for a given decoding complexity.

A. Minimizing the fields size

In order to maximize the approximate lower bound for the remaining entropy according to Theorem 7, we set $q = r$; this minimization of the finite field has no influence on the security against existing attacks. Furthermore, since $n > \psi \gg tV_d$, we have $\binom{n/V_d}{t}/\binom{\psi/V_d}{t} \approx (n/\psi)^t$. In general, we cannot assume $r \gg t$; therefore, we use the approximation

$$\binom{r}{t} \approx r^t \left(1 - \frac{t-1}{2r}\right)^t /(t!),$$

which is much tighter than $\binom{r}{t} \approx r^t/(t!)$. With Stirling's approximation for $t!$, this results in the estimate

$$E \lessapprox k\log r - t\log\left(\frac{nt}{\mathrm{e}\psi}\left(1 - \frac{t-1}{2r}\right)\right) - \frac{1}{2}\log(2\pi t) - 2. \quad (15)$$

B. Selecting the minimum distance for minutiae

In (15), the remaining entropy is independent of the minimum distance d enforced for minutiae and chaff points. However, the parameter d limits the maximum r to approximately $1 \leq r \leq 0.45n/V_{\lceil d/2 \rceil}$, where the factor

0.45 represents the maximum density of a random sphere packing [5].

On the other hand, d should not be smaller than the tolerance parameter δ used for minutiae matching, to limit false matchings of minutiae in the query fingerprint with chaff points during authentication. Setting $d = 2\delta$ will already completely prevent such false matchings with minutiae that are also present in T, but smaller values might already reduce their number to a minimum. According to [10], setting $\delta \approx (3/2)d$ is a good compromise. In the following, we base our analysis on this choice for d and will use $0.45n/V_{\lceil (3/4)\delta \rceil}$ as maximum value for r.

C. Optimizing the degree of the polynomial

The parameter k must be set, so that with sufficient probability the secret polynomial can be recovered efficiently from a genuine query fingerprint. Subsequently we analyze the expected complexity of this task. As in Section IV-E1, we denote the number of *correct matches*, i.e., the matches between the query fingerprint and the genuine minutiae, with m_c, and the number of *false matches*, i.e., the matches between the query fingerprint and the chaff points, with m_f. From Section III-B3 we know that decoding is only possible if $m_\mathrm{c} \geq m_\mathrm{f} + k$.

It has been shown in [5] that, on average, the Reed-Solomon decoding of the polynomial using ℓ points requires

$$\binom{m_\mathrm{c}+m_\mathrm{f}}{\ell} \left(\sum_{i=\max(\lceil \frac{\ell+k}{2} \rceil, \ell-m_\mathrm{f})}^{\min(\ell,m_\mathrm{c})} \binom{m_\mathrm{f}}{\ell-i} \binom{m_\mathrm{c}}{i} \right)^{-1}$$

trials, where the parameter ℓ must fulfill $k \leq \ell \leq \min(2m_\mathrm{c} - k, m_\mathrm{c} + m_\mathrm{f})$. This expression is difficult to analyze theoretically. Numerical evaluation shows that for $m_\mathrm{c} - k \leq m_\mathrm{f} \leq m_\mathrm{c} + 2m_\mathrm{c}/(m_\mathrm{c} - k)$, the decoding complexity is minimized for $\ell = 2m_\mathrm{c} - k$. In this case, the sum collapses to the term for $i = m_\mathrm{c}$ and hence the minimum decoding complexity is

$$C_{\min}(m_\mathrm{c}, m_\mathrm{f}, k) = \binom{m_\mathrm{c}+m_\mathrm{f}}{2m_\mathrm{c}-k} \binom{m_\mathrm{f}}{m_\mathrm{c}-k}^{-1}. \qquad (16)$$

In the case $m_\mathrm{f} = m_\mathrm{c} - k$, we have $\ell = 2m_\mathrm{c} - k = m_\mathrm{c} + m_\mathrm{f}$ and $C_{\min}(m_\mathrm{c}, m_\mathrm{f}, k)$ evaluates to 1. For $m_\mathrm{f} = m_\mathrm{c} - k + i$ with $i = 1, 2, \ldots, m_\mathrm{c}/(m_\mathrm{c} - k) - 1$ equation (16) yields

$$C_{\min}(m_\mathrm{c}, m_\mathrm{f}, k) = \frac{(2m_\mathrm{c} - k + 1) \cdots (2m_\mathrm{c} - k + i)}{(m_\mathrm{c} - k + 1) \cdots (m_\mathrm{c} - k + i)}.$$

This equation shows that, for $m_\mathrm{c} - k \leq m_\mathrm{f} < m_\mathrm{c} - k + m_\mathrm{c}/(m_\mathrm{c} - k)$, the minimum decoding complexity increases exponentially with $i = m_\mathrm{f} - m_\mathrm{c} + k < m_\mathrm{c}/(m_\mathrm{c} - k)$. Numerical evaluation reveals that the exponential growth continues (with slowing pace) for $m_\mathrm{f} - m_\mathrm{c} + k \geq m_\mathrm{c}/(m_\mathrm{c} - k)$. Consequently, we find that the decoding complexity is an exponential function in $m_\mathrm{f} - m_\mathrm{c} + k$.

On the other hand, the number m_c of correct matches will typically disperse considerably between different authentications due to variations in the fingerprint image quality. Thus, if k is larger than the expectation of $m_\mathrm{c} - m_\mathrm{f}$, the fraction of cases, in which decoding is not feasible anymore, can become quite high. As a consequence, we set k to the expectation of $m_\mathrm{c} - m_\mathrm{f}$ in order to optimize the remaining entropy while limiting the decoding complexity.

Depending on the specific distribution of the number of correct matches and the requirements on decoding complexity imposed by the application scenario, it may be appropriate to select smaller or larger values for k. For instance, if the False Reject Rate (FRR) observed for a certain k is too high, k must be decreased until the FRR becomes acceptable. On the other hand, if the FRR is very low, k could be carefully increased. We will investigate the impact of increasing or decreasing k in our numerical evaluation in Section VI-C.

We estimate the mean values for m_c and m_f as follows:

- It is reasonable to assume that the average number of correct matches is a linear function of t, i.e., $m_\mathrm{c} = \mu t$, where μ is the average match rate independent of t.
- If $rV_\delta \ll n$, the number of points in $\mathcal{M}$ covered by the tolerance areas $B_\delta(\mathbf{m}_i)$ around the chaff points $\mathbf{m}_i$ can be estimated as $(r - t)V_\delta$. (Since minutiae of the query fingerprint that lie within the tolerance area of a chaff points can still be correctly matched with a minutiae in T, this estimate is even conservative.) Therefore, we can estimate the average number m_f of false matches by $sf(r - t)V_\delta/n$, where s is the average number of surplus minutiae per query fingerprint, i.e., the average number of minutiae in the query fingerprints that do not match with the stored minutiae, and f is the number of fingers used.

Remark: As the surplus minutiae are those *not matching with genuine minutiae*, their number depends on the match rate. Precisely, we could estimate the number s of surplus minutiae per finger as $s \approx w - \mu t$, where w is the average number of (all) minutiae per query fingerprint. However, this would result in a term t^2 in the estimation of E, which would render analytical determination of the maximum achievable entropy much more difficult. Furthermore, the number w of minutiae per finger is also not constant but depends on the feature extraction algorithm used and quality filtering applied, and hence, we would end up with the same number of variable parameters in our results.

As we set k to the expectation of $m_c - m_f$, these estimations yield

$$k = t\mu - (r-t)\frac{sfV_\delta}{n}. \quad (17)$$

Using approximation (15) this yields $E \lessapprox f(t,r)$ with

$$f(t,r) = \left(t\mu - (r-t)\frac{sfV_\delta}{n}\right)\log r - t\log\left(\frac{nt}{e\psi}\left(1-\frac{t-1}{2r}\right)\right) - \frac{1}{2}\log(2\pi t) - 2.$$

We also use (17) to eliminate parameter k from the estimations (5) and (14) for the workload W of the polynomial reconstruction attacks, which allows numerical optimization of t and r with respect to practical attacks in Section VI-D.

D. Maximizing the Bound for the Entropy

For fixed δ, n, μ, s and f, we try to estimate the maximum remaining entropy E by finding the maximum of the function $f(t,r)$ over r. The maximum is assumed, where the first derivation $\frac{\partial f(t,r)}{\partial r}$ is zero. It is easy to see that this is equivalent to $t^2 + a(r)t + b(r) = 0$ with $a(r) = 2\mu nr + sfV_\delta r(3+\ln(r))$ and $b(r) = -2sfV_\delta r^2(\ln(r)+1)$. For $r > 0$, one of the two solutions is negative and can thus be neglected. Consequently, for every r, $f(t,r)$ takes its maximum at

$$t_0(r) = -a(r)/2 + \sqrt{a(r)^2/4 - b(r)}.$$

Consequently, the function $f(t_0(r),r)$ upper bounds E for a given r, and the maximum of $f(t_0(r),r)$ over r yields a general upper bound for E. Thus, we can estimate the best provable security bound according to Theorems 1 and 4 that can be achieved for given δ, n, μ, s and f, by numerically determining the maximum of $f(t_0(r),r)$ over the relevant range of r. As argued in Section V-A, it is reasonable to set $d = \lceil (3/2)\delta \rceil$; hence, the relevant range is given by $1 \le r \le 0.45n/V_{\lceil (3/4)\delta \rceil}$ (see Section V-B), where the factor 0.45 represents the density of a random sphere packing [5].

Since for fixed $t, r \ge 1$, the value $f(t,r)$ is monotonically increasing with the match rate μ, we can determine the minimum value $\mu_{\min}$, for which the maximum of $f(t_0(r),r)$ is greater than a certain security level S. Since $E \lessapprox f(t_0(r),r)$, this value $\mu_{\min}$ is an approximate lower bound for the average match rate required to obtain a scheme with security 2^S according to Theorem 1 and Theorem 7, so that in the average case the polynomial can be recovered with one trial.

VI. Results

We evaluate whether and to what extent a (heuristically) provably secure fuzzy fingerprint vault is feasible. In particular, for different values for δ and for typical values for n, ψ and s we determine the minimum match rates required to achieve a security of 2^{50} according to Theorem 1 and Theorem 7. We compare these minimum match rates with match rates observed in practice.

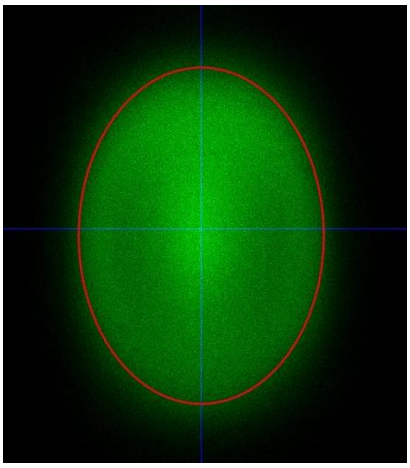


Figure 2. Spatial distribution of minutiae extracted with feature extractor MINDTCT [47] from 82800 fingerprints and the ellipse $\mathcal{E}$ from where minutiae are considered. The brightness of pixels corresponds to the frequency of minutiae occurrence at this position.

A. Evaluation of Minutiae Distribution

In order to estimate n and ψ, we have empirically determined the spatial distribution of minutiae within the fingerprint image. We evaluated the location of 5.8 million minutiae extracted with NIST's MINDTCT feature extraction algorithm [47] from 82800 imprints that were taken from 9200 fingers with 3 different sensors having 500 DPI. The fingerprints were taken from a non-public database set up in the course of a previous project of the German Federal Office for Information Security (Bundesamt für Sicherheit in der Informationstechnik). For this evaluation, the fingerprints had been pre-aligned, so that the center of mass of all minutiae coincides with the image center and the longest distance between two minutiae locations was vertically aligned.

It turned out that 83% of all minutiae occurred in an area defined by an ellipse that covers approximately 87000 pixels, which roughly corresponds to 2.25 cm^2. Outside this ellipse, the density of minutiae decreases drastically. Therefore, it is reasonable to restrict the fuzzy vault to minutiae and chaff points inside this area. This gives an estimate $n \approx 87000 \cdot f$, where f is the number of fingers, from which the minutiae are gathered. The distribution of the minutiae positions and the ellipse are shown in Figure 2. This yields $fs/n \approx s/87000$, which makes our estimation $E \lessapprox f(t,r)$ for the maximum achievable entropy bound E independent of the number f of fingers.

Table I
MINIMUM MATCH RATES REQUIRED TO ACHIEVE A PROVABLE SECURITY OF 50 BITS.

	$\delta = 5$	$\delta = 7$	$\delta = 10$
$s = 20$	82.2%	89.6%	97.0%
$s = 35$	87.9%	95.5%	-
$s = 50$	91.7%	99,1%	-

Table II
LINEAR FACTOR, BY WHICH THE MINIMUM MATCH RATES GIVEN IN TABLE I DECREASE WITH INCREASING k.

	$\delta = 5$	$\delta = 7$	$\delta = 10$
$s = 20$	0.67%	0.54%	0.41%
$s = 35$	0.54%	0.38%	-
$s = 50$	0.45%	0.28%	-

The maximum frequency of a minutiae location was 112, which corresponds to a maximum probability of a minutiae location inside the ellipse of approximately $112/5800000/0.83 \approx 2^{-15.4}$. This results in an approximation $n/\psi \approx 2$. This approximation is independent from the number f of fingers used for the fuzzy vault, as both ψ and n scale linearly with f.

We stress that our estimate is valid for minutiae extracted with the MINDTCT algorithm. As shown in [54], other feature extractors exhibit considerably different minutiae placement density functions, and thus, the maximum probability of a minutiae location may differ as well.

B. Estimating the number of surplus minutiae

According to [4], a good-quality live-scan fingerprint has 20–70 minutiae. Since $f(t, r)$ decreases with an increasing average number s of minutiae per query fingerprint not matching with genuine minutiae, it might be a good idea to use only the most reliable minutiae of the query fingerprints, e.g., by evaluating minutiae quality indices output by the feature extraction algorithm. However, the extent of the filtering should be carefully balanced with the match rates achieved with the reduced number of minutiae. We will subsequently consider the range $20 \leq s \leq 50$.

C. Numerical Parameter Optimization for Provable Security

In the previous sections, we found the approximations $n/\psi \approx 2$ and $fs/n \approx s/87000$ from empirical data. Using these estimations and various values for δ and s, we applied the method described in Section V-D to determine the minimum match rate required to achieve a security level of 2^{50} according to Theorem 1 and Theorem 7. We numerically computed the maximum value of the function $f(t_0(r), r)$ over the range $1 \leq r \leq 0.45n/V_{\lceil (3/4)\delta \rceil}$ (see Section V-B) with the maximum value $n = 10 \cdot 87000$, i.e., for maximum range of chaff points possible for 10 fingers, using the computer algebra system PARI/GP. The minimum match rates, at which this maximum exceeds 2^{50}, are listed in Table I for different values of δ and s. A "–" denotes that a remaining entropy of 50 is not achieved at all.

The security bounds are very sensitive to changes of the match rate. For instance, for the parameters given in Table I, a decrease of the match rate by only 2% results in a reduction of the achievable security of 12 to 38 bits; a larger reduction is observed for higher match rates.

As explained in Section V-C, under specific circumstances it may be reasonable to select k greater than our choice $k_0 := t\mu - (r - t)sfV_\delta/n$, particularly if the dispersion of the number of correct matches is small, or if a larger decoding complexity is acceptable. On the other hand, if the False Reject Rate (FRR) observed for $k = k_0$ is too high or the decoding of the polynomial takes too much time, k must be decreased. Any decrease of k from the assumed optimal value k_0 results in an increase of the minimum match rate required for a certain security level, and any increase of k results in a decrease of the minimum match rate. In particular, setting $k = k_0 + \epsilon$ with $\epsilon > 0$ increases the entropy estimation 15 by $\epsilon \log(r)$. For a given match rate μ, this results in the same amount of entropy as setting $k = k_0$ with match rate $\mu + \epsilon/t$. Thus, for a given security level, decreasing k by ϵ compensates an decrease of the match rate by ϵ/t. As a consequence, the minimum match rates required for a security level of 2^{50} with $k = k_0 + \epsilon$ can be estimated by subtracting $\epsilon/t_{\max}$ from the values given in Table I, where $t_{\max}$ is the value of $t_0(r)$, for which $f(t_0(r), r)$ is maximal. Analogously, the minimum match rates with $k = k_0 - \epsilon$ can be estimated by adding $\epsilon/t_{\max}$ to the values given in Table I. We give the respective values of $1/t_{\max}$ in Table II.

We give an example how Table II can be used: According to Table I, for $\delta = 5$ and $s = 35$, at least a match rate $\mu = 87.9\%$ is required to achieve a security of 2^{50}, given that we set $t\mu - (r - t)sfV_\delta/n$ and select r and $t = t_0(r)$ so that $f(t, r)$ is maximized. However, if the False Reject Rate (FRR) observed for these parameters is too high or the decoding of the polynomial takes too much time, k must be decreased. If setting $k = t\mu - (r - t)sfV_\delta/n - 3$ results in an acceptable FRR and decoding performance, we have $\epsilon = 3$. This decrease of k implies that the match rate has to be at least $\mu' = 87.9\% + 3 \cdot 0.54\% = 89.52\%$ to achieve the desired security of 2^{50}, where the value 0.54 is taken from Table II.

To get a feeling for the number of minutiae and thus for the number of fingers needed for a provable secure scheme, we evaluate the minimum value t, for which we still obtain a remaining entropy of 2^{50} for a given μ. For this evaluation we apply the following method.

First, we observe that $t_0(r)$ is continuous and unbounded for $r > 0$ and is zero for $1/\mathrm{e}$. Thus, for every $t' > 0$ there is a r' with $t' = t_0(r')$; by definition of $t_0(r)$, this pair (t', r') maximizes the function $f(t, r')$ over t. Consequently, it suffices to search through all pairs $(t_0(r), r)$ to find the minimal t with $f(t, r) \geq 2^{50}$.

On the other hand, the approximation of the remaining entropy E by the continuous function $f(t, r)$ will result in an artificially smooth curve for the minimal t. In particular, in the definition of f we have replaced k by a real number, whereas in practice, k can only take integer values. The small deviations of the truncated integer k from its real valued optimum imply a corresponding deviation of the achievable security E and hence, of the minimal t required for a certain value of E. To obtain a more realistic estimation of the minimal t, we set $k_0(r) = \lfloor t_0(r)\mu - (r - t_0(r))sfV_\delta/n \rfloor$ and determine the minimal $t_0(r)$ for that (15) yields at least a value of $E \geq 2^{50}$ with $t = t_0(r)$ and $k = k_0(r)$.

Figure 3 shows the minimal number t of minutiae required for a security of 2^{50} as a function of the average match rate μ for various parameters δ and s.

These curves also allow estimating the impact of selecting a larger k to the minimum value t of minutiae. As explained above, selecting $k = k_0(r) + \epsilon$ compensates a decrease of the match rate by ϵ/t, and analogously, choosing $k = k_0(r) - \epsilon$ equates an increase of the match rate by ϵ/t. Therefore, for small ϵ, the minimum value of t yielding a security of 2^{50} with $k = k_0(r) \pm \epsilon$ and a match rate μ can be estimated as the value of t corresponding to $\mu \mp \epsilon/t_0$ in Figure 3, where t_0 is the value of t indicated in Figure 3 for μ.

We give an example: for $\delta = 5$ and $s = 35$, a match rate $\mu = 0.9$ requires at least $t = 68$ minutiae in the template. If for this t, the corresponding optimal r (maximizing function $f(t, r)$, i.e., the r with $t = t_0(r)$) and $k_0 = \lfloor t\mu - (r - t)sfV_\delta/n \rfloor$, the False Reject Rate (FRR) observed is too high, k must be decreased until the FRR becomes acceptable. If setting $k = \lfloor t\mu - (r-t)sfV_\delta/n \rfloor - 3$ results in an acceptable FRR, we have $\epsilon = 3$ and, using Table II, obtain a minimum match rate of $\mu' = 90\% + 3 \cdot 0.54\% = 91.62\%$. For this value, we get from Figure 3 that only a minimum of $t = 58$ minutiae are required. We have taken the exact values from our evaluation data, from which the curves in Figure 3 have been drawn.

D. Numerical Parameter Optimization for Practical Security

In this section, we determine the minimal number of minutiae required for a given match rate μ and fixed parameters δ and s, for which a security of 2^{66} can be achieved against existing attacks. We do this by numerically evaluating the estimates from (5) and Theorem 9 for the expected number of operations needed for polynomial reconstruction. Precisely, for each t we maximize the estimates for W according to (5) and Theorem 9, respectively, with respect to r over the relevant range $t + 1 \leq r \leq 0.45n/V_{\lceil (3/4)\delta \rceil}$ (see Section V-B), and identify the minimal t, for which an r from this range exists so that the respective security exceeds 2^{66}. Again, we deploy the computer algebra program PARI/GP.

Figure 4 shows the dependency of the minimal number of minutiae in the template required to achieve a security of 2^{66} against the polynomial non-optimized attack proposed in [12] according to (5). This estimation (5) also provides an upper bound for the workload of the smart polynomial reconstruction attack, and hence, the minimum number of minutiae indicated in Figure 4 provides a lower bound for the minimum number of minutiae needed to ensure the same security level of 2^{66} with respect to the smart polynomial reconstruction, which we consider as the best known attack. An approximate lower bound of the security against this attack is given by Theorem 9, and consequently, this estimation can be used to determine an upper bound for the number of minutiae needed to achieve a certain security level. Figure 5 shows the dependency of the minimal number of minutiae in the template required to achieve a security of 2^{66} against the smart polynomial reconstruction attack based on the estimation of Theorem 9. Since our empirical results in [10] indicates that the number of minutiae required according to Figure 4 can only be obtained by used $f \geq 2$ fingers per person, and as the estimate of Theorem 9 decreases with f, we will assume $f = 2$ (the estimate (5) is independent of f).

For all considered parameters, the optimal value for k is between 18 and 49. This implies that the number $117k \log^2(k)$ of operations needed for a polynomial interpolation (see Section IV-E2) is between $2^{15.1}$ and $2^{17.5}$, and thus, the 2^{66} operations used as security level roughly correspond to 2^{50} trials. As a consequence, our security bound considered for practical attacks is comparable to the security bound used for provable security in Section VI-C.

An interesting observation is that increasing the number of chaff points does not generally increase security: for each t, there is an optimal value for parameter r, for which the estimates given by (5) and Theorem 9 are maximized. If r is increased further, the estimated workload of the polynomial reconstruction attack for optimally chosen k, decreases. This observation can be explained by the influence of r on the number of false matches. Specifically, the average number of false matches increases linearly with the number $(r-t)$ of chaff points. This increased number of false matches requires a smaller value for parameter k in order to ensure efficient

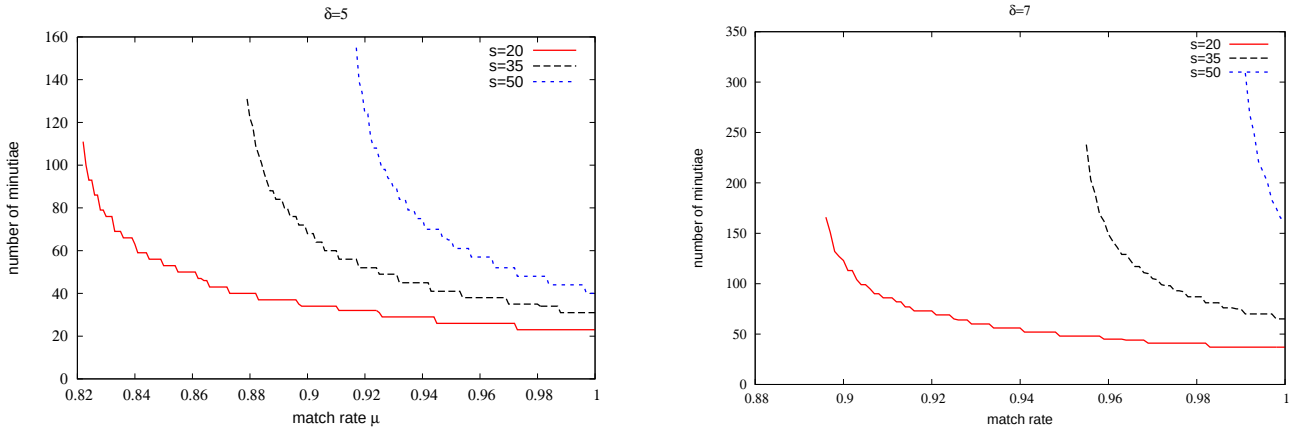


Figure 3. Dependency of the minimal number t of minutiae on the average match rate μ for a security of $E \geq 2^{50}$, for (a) $\delta = 5$ and (b) $\delta = 7$, respectively, and different numbers s of surplus minutiae per finger.

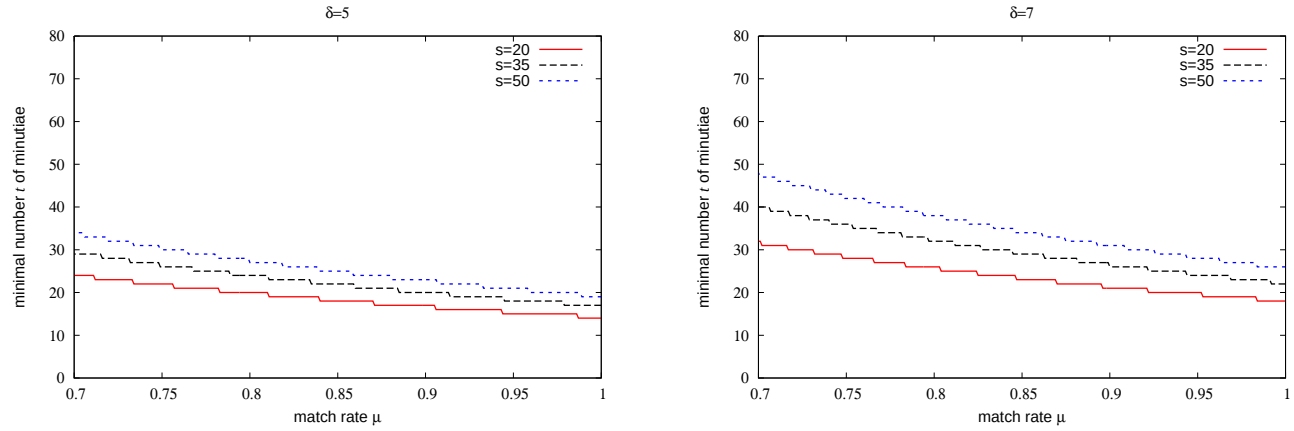


Figure 4. Dependency of the minimal number t of minutiae on the average match rate μ for $\delta = 5$ and a security of 2^{66} against the non-optimized polynomial reconstruction attack of [12] for (a) $\delta = 5$ and (b) $\delta = 7$ and for different numbers s of surplus minutiae per finger.

decoding of the polynomial, and a smaller k reduces the workload of the attack significantly.

Our results allow a critical review of our assumptions $r \ll 0.45n/V_{\lceil d/2 \rceil}$ and $r \ll n/V_d$ used in Section IV. For all parameters μ and s considered, the optimal r fulfills $r < 408$ for $\delta = 5$ and $r < 312$ for $\delta = 5$. This implies that, for $\delta = 5$ and $f \geq 2$, the optimal r is by a factor 4 smaller than the maximum value $0.45n/V_{\lceil d/2 \rceil}$ with $d \approx (3/2)\delta$ (see Section V-B), and for $\delta = 7$ and $f \geq 3$, it is by a factor 3.5 smaller than the maximum value. Consequently, at least for $\delta = 5$ and $f \geq 2$, or for $\delta = 7$ and $f \geq 3$, respectively, our assumption that the attack method of [11] is not very efficient is justified. Unfortunately, the validity of our assumption $rV_d \ll n$ used for the proof of Lemma 8 is less clear: for $\delta = 5$ and $f = 2$ as well as for $\delta = 7$ and $f = 3$, the fraction $n/(rV_d)$ is approximately 2.2. Thus, unless r is chosen considerably smaller than its optimum, which is possible by increasing the number t of minutiae, we must expect some inaccuracy in our approximation that, on average, each selected chaff point reduces the number of choices for the subsequent points by V_d. This inaccuracy propagates to the estimation of Theorem 9. On the other hand, our upper bound $1/\psi$ for the probability of a minutiae occurring at location $\mathbf{m}_{j_i}$ for all minutiae $\mathbf{m}_{j_i}$ processed by the smart polynomial reconstruction attack is very conservative, because from Figure 2 we can see that the area with the highest frequency of minutiae occurrence, which is the bright area in the center, is quite small. Therefore, we are very confident that Theorem 9 still underestimates the number of trials needed for the smart polynomial reconstruction attack.

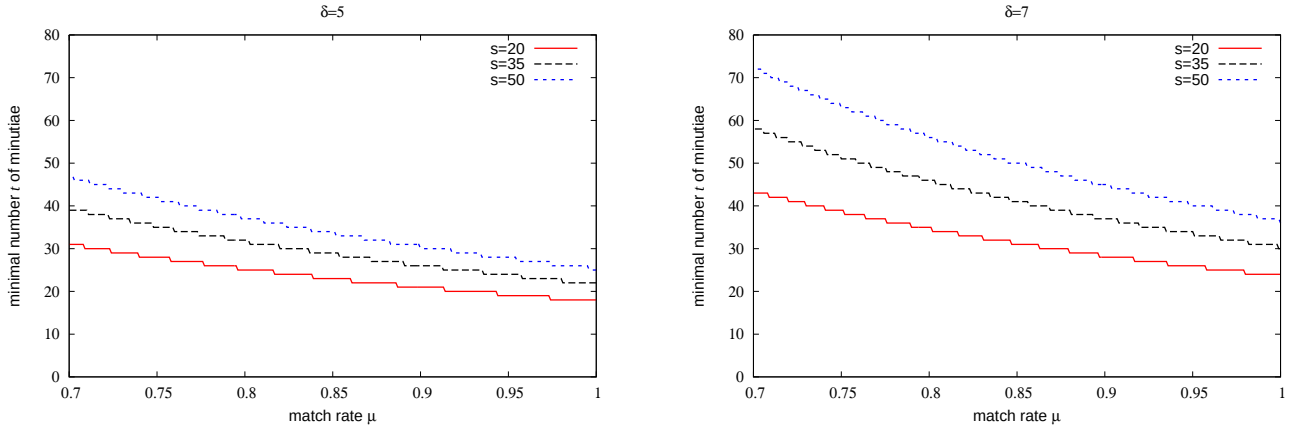


Figure 5. Dependency of the minimal number t of minutiae on the average match rate μ for $\delta = 5$ and a security of 2^{66} against the smart polynomial reconstruction attack presented in Section IV-E2 for (a) $\delta = 5$ and (b) $\delta = 7$ and for different numbers s of surplus minutiae per finger.

E. Comparison with Empirical Data

The minimum match rates required for provable security are quite high. According to [4], matchings conducted by a human expert results in rates of approximately 90%. Automatic matching algorithms only operating on the minutiae data will yield considerably lower rates, depending on the method, by which minutiae correspondences are identified and the matching tolerance applied. For instance, the distribution of the distance of matching minutiae reported in literature (see [55]) implies that a tolerance (with respect to Euclidean distance) of $\delta < 10$ will significantly reduce the match rates. The presence of chaff points will further reduce the performance of minutiae matcher algorithms.

On the other hand, the match rates can be greatly enhanced by using several minutiae measurements (per finger) during enrollment to minimize measurement noise. For instance, in [10], we present an implementation of the fuzzy fingerprint vault, which applies several minutiae measurements per finger during enrollment and uses only those minutiae that have been detected in all measurements. Furthermore, the minutiae locations are set to the mean value over the measurements. Empirical evaluations show that this considerably increases the average match rates, but, at the same, reduces the number of minutiae available per finger. Hence, several fingers must be used to achieve the minimum values for t indicated in Figure 3. Furthermore, the dispersion found in the match rates is large so that, in order to limit the False Rejection Rate, is seems necessary to choose k smaller than the expectation of $m_c - m_f$ (see Section V-C), which further increases the required minimum match rates as shown in Table II.

We found minutiae quality filtering during authentication using quality indices provided by the feature extractor to be quite effective to reduce the number s of surplus (non-matching) minutiae per query fingerprint and, consequently, the number of false matches. However, the filtering should not exceed a certain extent in order to avoid disproportional reduction of the match rate. Furthermore, the false match rates observed were 20% to 60% higher than our estimation $sf(r-t)V_\delta/n$ in Section V-C predicts. This effect is presumably due to failures in the alignment of the query fingerprint to the vault.

Based on the observed statistics on match rates and number of reliable minutiae, and given the results in Table I, we conclude that provable security seems out of reach, unless the average number of surplus minutiae per query fingerprint can be further reduced by improved quality filtering methods. Details on the empirical data and our interpretation are given in [10].

On the other hand, our evaluation in [10] shows that strong security (comparable to 64 bit keys) against the (non-optimized) polynomial reconstruction attack can be achieved using 2 fingers per individual. Considering our improvement by the smart polynomial reconstruction, an additional security margin should be added. Based on the data provided in [10] we can estimate that the same security can be achieved against our optimized smart polynomial reconstruction attack using 3 fingers per user. As explained in Section IV-E, the number of chaff point should not be too close to the maximum possible to render the attack method described in [11] inefficient, and the FAR should be determined for the chosen parameters to allow estimation of the effort for a fingerprint dictionary attack.

VII. Conclusions

Our analysis shows that a provably secure fuzzy fingerprint vault can hardly be achieved in practice. The required rate of minutiae in the vault matching with those in the query fingerprints so high that it seems only achievable by powerful quality filtering during enrollment. However, this filtering approach conflicts with the requirement for a large number of minutiae in the vault. Given the empirical data on match rates in the literature, in particular our analysis in [10], provable security seems out of reach.

The usage of minutiae orientations as additional discriminating data could surely increase the information content of the templates. However, minutiae directions bear strong dependencies with their spatial location and with directions of nearby minutiae: according to [45], "minutiae points in different regions of the fingerprint domain are observed to be associated with different region-specific minutiae directions", and "minutiae points that are spatially close tend to have similar directions with each other". Consequently, a template using both spatial location and orientation of minutiae contains considerable redundancy and makes an analysis of the entropy of the feature vector very difficult. Moreover, since the estimated entropy loss in the security bounds in Section IV-C increases linearly with the number of bits in the template, this decreases the percentage of entropy that actually contributes to the provable security estimates.

On the other hand, our investigation of the most efficient attack methods indicates that the theoretical lower bounds for security are far from being tight. The underlying computational problem (polynomial reconstruction) is believed to be hard and has been repeatedly proposed as a basis for the security of cryptographic techniques [52]. As a consequence, the match rates and number of minutiae required to achieve security against the existing attacks are much lower than the numbers for provable security. Still, the empirical data presented in [10] show that at least two fingers per user must be used to achieve a level of security equivalent to a 50 bit cryptographic key. However, we stress that there is no evidence that our optimized polynomial reconstruction method is indeed the most efficient attack. In particular, before an implementation of the fuzzy fingerprint vault can be claimed to be secure, it must be verified that the False Accept Rate (FAR) is in a range that ensures that a fingerprint dictionary attack is inefficient. Unfortunately, this requires a very large number of impostor matches, precisely, in the range of 1/FAR. The numbers of fingers used in existing investigations for the fuzzy fingerprint vault, in particular in [6], [7], [8], [9] and [10], were far too small to assess if an adequate security level against the fingerprint dictionary attack can be achieved.

Summarizing, our results seem to indicate that although a provable secure fuzzy fingerprint vault is out of reach, it can provide sufficient security against practical attacks if several fingers are used.

Finally, secure biometric template protection schemes may also be achievable using completely different constructions. For instance, there exist approaches to apply the fuzzy commitment scheme to fingerprints. As shown in [23], the entropy loss in the fuzzy commitment is much lower than in the fuzzy vault. However, since the fuzzy commitment scheme only tolerates errors that are small with respect to the Hamming metric [13], sophisticated encoding and signal processing techniques must be applied to compensate spatial rotations and translations of the fingerprint, as well as permutations, detections and insertions of the detected minutiae. Several promising techniques have been proposed, in particular, usage of fingerprint ridge patterns as biometric feature [20][56], transformation of the minutiae data to the frequency domain [57] and using the characteristic vector of minutiae occurrence with respect to a grid [58][59]. However, we are not aware of any comprehensive security analysis for these approaches based on estimations for the feature vector's entropy and the error correction required without manual alignment of the fingerprints.

Acknowledgments

This work was conducted as part of the project "BioKeyS Pilot-DB" of the German Federal Office for Information Security (Bundesamt für Sicherheit in der Informationstechnik).

We would like to thank Patrick Jäger for his assistance in the numerical calculations and visualization.

References

[1] J. Merkle, M. Niesing, M. Schwaiger, H. Ihmor, and U. Korte, "Provable security for the fuzzy fingerprint vault," in *Proceedings of the 5th International Conference Internet Monitoring and Protection (ICIMP 2010)*, J. L. Mauri and M. Popescu, Eds. IEEE Computer Society, 2010, pp. 65–73.

[2] J. Breebaart, C. Busch, J. Grave, and E. Kindt, "A reference architecture for biometric template protection based on pseudo identities," in *BIOSIG 2008: Biometrics and Electronic Signatures*, ser. Lecture Notes in Informatics, A. Brömme and C. Busch, Eds., vol. 137. Gesellschaft für Informatik, 2008, pp. 25–38.

[3] A. Juels and M. Sudan, "A fuzzy vault scheme," in *Proceedings of the 2002 IEEE International Symposium on Information Theory*. IEEE, 2002, p. 408.

[4] U. Uludag, S. Pankanti, S. Prabhakar, and A. Jain, "Biometric cryptosystems: Issues and challenges," *Proceedings of the IEEE*, vol. 92, no. 6, pp. 948–960, 2004.

[5] C. Clancy, N. Kiyavash, and D. Lin, "Secure smartcard-based fingerprint authentication," in *Proceedings of the 2003 ACM SIGMM Workshop on Biometrics Methods and Applications (WBMA '03)*. ACM, 2003, pp. 45–52.

[6] U. Uludag, S. Pankanti, and A. Jain, "Fuzzy vault for fingerprints." in *Proceedings of the 5th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, ser. Lecture Notes in Computer Science, vol. 3546. Springer, 2005, pp. 310–319.

[7] U. Uludag and A. Jain, "Securing fingerprint template: Fuzzy vault with helper data," in *Proceedings of the 2006 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2006), Workshop on Privacy Research In Vision*, C. Schmid, S. Soatto, and C. Tomasi, Eds. IEEE Computer Society, 2006, pp. 163–169.

[8] K. Nandakumar, A. Jain, and S. Pankanti, "Fingerprint-based fuzzy vault: Implementation and performance." *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 4, pp. 744–757, 2007.

[9] J. Li, X. Yang, J. Tian, P. Shi, and P. Li, "Topological structure-based alignment for fingerprint fuzzy vault," in *Proceedings of the 19th International Conference on Pattern Recognition (ICPR '08)*. Omnipress, 2008, pp. 1–4.

[10] J. Merkle, M. Niesing, M. Schwaiger, H. Ihmor, and U. Korte, "Performance the fuzzy vault for multiple fingerprints (short version)," in *BIOSIG 2010: Biometrics and Electronic Signatures*, ser. Lecture Notes in Informatics, A. Brömme and C. Busch, Eds., vol. P-164. Gesellschaft für Informatik, 2010, pp. 57–72.

[11] E.-C. Chang, R. Shen, and F. W. Teo, "Finding the original point set hidden among chaff," in *Proceedings of the 2006 ACM Symposium on Information, Computer & Communication Security (ASIACCS)*. ACM, 2006, pp. 182–188.

[12] P. Mihailescu, A. Munk, and B. Tams, "The fuzzy vault for fingerprints is vulnerable to brute force attack," in *BIOSIG 2009: Biometrics and Electronic Signatures*, ser. Lecture Notes in Informatics, A. Brömme and C. Busch, Eds., vol. 155. Gesellschaft für Informatik, 2009, pp. 43–54.

[13] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data." *SIAM Journal on Computing*, vol. 38, no. 1, pp. 97–139, 2008.

[14] C. Soutar, D. Roberge, A. Stoianov, R. Gilroy, and B. V. Kumar, "Biometric encryption using image processing," in *Proceedings of the SPIE, Optical Security and Counterfeit Deterrence Techniques II*, R. L. van Renesse, Ed., vol. 3314, 1998, p. 178Ű188.

[15] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy of biometric-based authentication systems," *IBM Systems Journal*, vol. 40, 2001.

[16] J.-P. Linnartz and P. Tuyls, "New shielding functions to enhance privacy and prevent misuse of biometric templates," in *Proceedings of the 4th international conference on Audio- and video-based biometric person authentication (AVBPA'03)*. Springer, 2003, pp. 393–402.

[17] Y. Chang, W. Zhang, and T. Chen, "Biometrics-based cryptographic key generation," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '04)*. IEEE Computer Society, 2004, pp. 2203–2206.

[18] C. Chen, R. Veldhuis, T. Kevenaar, and A. Akkermans, "Multibits biometric string generation based on the likelyhood ratio," in *Proceedings of the IEEE Conference on Biometrics: Theory, Applications and Systems (BTAS '07)*. IEEE Computer Society, 2007, pp. 1–6.

[19] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *Proceedings of the 6th ACM Conference on Computer and Communication Security*. ACM, 1999, pp. 28–36.

[20] U. Martini and S. Beinlich, "Virtual PIN: Biometric encryption using coding theory," in *BIOSIG 2003: Biometrics and Electronic Signatures*, ser. Lecture Notes in Informatics, A. Brömme and C. Busch, Eds., vol. 31. Gesellschaft für Informatik, 2003, pp. 91–99.

[21] T. A. M. Kevenaar, G. J. Schrijen, M. van der Veen, A. H. M. Akkermans, and F. Zuo, "Face recognition with renewable and privacy preserving binary templates," in *Proceedings of the Fourth IEEE Workshop on Automatic Identification Advanced Technologies*. IEEE Computer Society, 2005, pp. 21–26.

[22] F. Hao, R. Anderson, and J. Daugman, "Combining crypto with biometrics effectively," *IEEE Transaction on Computers*, vol. 55, pp. 1081–1088, 2006.

[23] U. Korte, M. Krawczak, J. Merkle, R. Plaga, M. Niesing, C. Tiemann, H. Vinck, and U. Martini, "A cryptographic biometric authentication system based on genetic fingerprints," in *Sicherheit 2008 - Sicherheit, Schutz und Zuverlässigkeit*, ser. Lecture Notes in Informatics, A. Alkassar and J. Siekmann, Eds., vol. 128. Gesellschaft für Informatik, 2008, pp. 263–276.

[24] Y. Lee, K. Bae, S. Lee, K. Park, and J. Kim, "Biometric key binding: Fuzzy vault based on iris images," in *Advances in Biometrics*, ser. Lecture Notes in Computer Science, S.-W. Lee and S. Li, Eds. Springer, 2007, vol. 4642, pp. 800–808.

[25] A. Juels and M. Sudan, "A fuzzy vault scheme," *Designs, Codes and Cryptography*, vol. 38, pp. 237–257, 2006.

[26] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," in *Advances in Cryptology - EUROCRYPT 2004*, ser. Lecture Notes in Computer Science, C. Cachin and J. Camenisch, Eds., vol. 3027. Springer-Verlag, 2004, pp. 523–540.

[27] S. Yang and I. Verbauwhede, "Automatic secure fingerprint verification system based on fuzzy vault scheme," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'05)*, 2005, pp. 609–612.

[28] P. Li, X. Yang, K. Cao, X. Tao, R. Wang, and J. Tian, "An alignment-free fingerprint cryptosystem based on fuzzy vault scheme," *Journal of Network and Computer Applications*, vol. 33, pp. 207–220, 2010.

[29] A. Nagar, K. Nandakumar, and A. K. Jain, "Securing fingerprint template: Fuzzy vault with minutiae descriptors," in *Proceedings of the 19th International Conference on Pattern Recognition (ICPR '08)*. Omnipress, 2008, pp. 1–4.

[30] W. J. Scheirer and T. E. Boult, "Cracking fuzzy vaults and biometric encryption," in *Proceedings of the Biometrics Symposium 2007*, 2007, pp. 1–6.

[31] A. Kholmatov and B. Yanikoglu, "Realization of correlation attack against the fuzzy vault scheme," in *Proceedings of the SPIE, Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, E. J. Delp, Ed., vol. 6819, 2008, pp. 68 190O–68 190O–7.

[32] K. Nandakumar, A. Nagar, and A. K. Jain, "Hardening fingerprint fuzzy vault using password," in *Proceedings of Advances in Biometrics, International Conference (ICB 2007)*, ser. Lecture Notes in Computer Science, S.-W. Lee and S. Z. Li, Eds., vol. 4642. Springer, 2007, pp. 927–937.

[33] J. Gu and J. Zhou, "A novel model for orientation field of fingerprints," in *Proceeding of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003)*. IEEE Computer Society, 2003, pp. 2–493.

[34] S. C. Dass, Y. Zhu, and A. K. Jain, "Statistical models for assessing the individuality of fingerprints," *IEEE Transactions on Information Forensics and Security*, vol. 2, pp. 391–401, 2007.

[35] R. Plaga, "Biometric keys: Suitable uses and achievable information content," *International Journal on Information Security*, vol. 8, no. 6, pp. 447–454, 2009.

[36] T. Kevenaar, U. Korte, J. Merkle, M. Niesing, H. Ihmor, C. Busch, and X. Zhou, "A reference framework for the privacy assessment of biometric encryption systems," in *BIOSIG 2010: : Biometrics and Electronic Signatures*, ser. Lecture Notes in Informatics, A. Brömme and C. Busch, Eds., vol. P-164. Gesellschaft für Informatik, 2010, pp. 45–56.

[37] R. Canetti, "Towards realizing random oracles: Hash functions that hide all partial information," in *Advances in Cryptology - CRYPTO '97*, ser. Lecture Notes in Computer Science, B. S. Kaliski, Ed., vol. 1294. Springer, 1997, pp. 455–469.

[38] M. Liu, X. Jiang, and A. C. Kot, "Fingerprint reference point detection," *EURASIP Journal on Applied Signal Processing*, vol. 2005, pp. 498–509, 2005.

[39] P. Tuyls and J. Goseling, "Capacity and examples of template-protecting biometric authentication systems," in *Biometric Authentication: ECCV 2004 International Workshop, BioAW 2004*. Springer, 2004, pp. 158–170.

[40] A. Bhargav-Spantzel, A. Squicciarini, and E. Bertino, "Privacy preserving multi-factor authentication with biometrics," in *DIM '06: Proceedings of the Second ACM Workshop on Digital Identity Management*. ACM, 2006, pp. 63–72.

[41] T. Ignatenko, "Secret-key rates and privacy leakage in biometric systems," Ph.D. dissertation, Eindhoven University of Technology, 2009.

[42] J. O. Pliam, "On the incomparability of entropy and marginal guesswork in brute-force attacks," in *Progress in Cryptology - INDOCRYPT 2000*, B. K. Roy and E. Okamoto, Eds. Springer, 2000, pp. 67–79.

[43] I. R. Buhan, "Cryptographic keys from noisy data, theory and applications," Ph.D. dissertation, University of Twente, 2008.

[44] S. Pankanti, S. Prabhakar, and A. Jain, "On the individuality of fingerprints," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001)*, vol. 1, p. 805, 2001.

[45] Y. Zhu, S. C. Dass, and A. Jain, "Statistical models for assessing the individuality of fingerprints." *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3-1, pp. 391–401, 2007.

[46] J. Chen and Y. S. Moon, "A statistical study on the fingerprint and minutiae distribution," in *Proceedings of the 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, P. Duhamel and L. Vandendorpe, Eds. IEEE Computer Society, 2006, pp. 169–172.

[47] C. Watson, M. Garris, E. Tabassi, C. Wilson, M. McCabe, S. Janet, and K. Ko, *User's Guide to NIST Biometric Image Software (NBIS)*, National Institute of Standards and Technology, 2007.

[48] D. A. Stoney, "Distribution of epidermal ridge minutiae," *American Journal of Physical Anthropology*, vol. 77, pp. 367–376, 1988.

[49] S. Z. Li and A. K. Jain, Eds., *Encyclopedia of Biometrics*. Springer US, 2009.

[50] A. Kiayias and M. Yung, "Cryptographic hardness based on the decoding of reed-solomon codes," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2752–2769, 2008.

[51] V. Guruswami and A. Vardy, "Maximum-likelihood decoding of reed-solomon codes is NP-hard," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2249–2256, 2005.

[52] A. Kiayias and M. Yung, "Directions in polynomial reconstruction based cryptography," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E87-A, no. 5, pp. 978–985, 2004.

[53] J. v. z. Gathen and J. Gerhard, *Modern Computer Algebra*, 2nd ed. Cambridge University Press, 2003.

[54] E. Tabassi, P. Grother, W. Salamon, and C. Watson, "Minutiae interoperability," in *BIOSIG 2009: Biometrics and Electronic Signatures*, ser. Lecture Notes in Informatics, A. Brömme and C. Busch, Eds., vol. P-155. Gesellschaft für Informatik, 2009, pp. 13–30.

[55] A. Jain, S. Prabhakar, and S. Pankanti, "On the individuality of fingerprints," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001)*. IEEE Computer Society, 2001, pp. I:805–812.

[56] P. Tuyls, A. Akkermans, T. Kevenaar, G. J. Schrijen, A. Bazen, and R. Veldhuis, "Practical biometric authentication with template protection." in *Proceedings of the 5th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, ser. Lecture Notes in Computer Science, T. Kanade, A. K. Jain, and N. K. Ratha, Eds., vol. 3546. Springer, 2005, pp. 436–446.

[57] H. Xu and R. N. Veldhuis, "Spectral minutiae representations for fingerprint recognition," in *Proceedings of the Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP 2010)*, I. Echizen, J.-S. Pan, D. Fellner, A. Nouak, A. Kuijper, and L. C. Jain, Eds. IEEE Computer Society, 2010, pp. 341–345.

[58] A. Arakala, J. Jeffers, and K. Horadam, "Fuzzy extractors for minutiae-based fingerprint authentication," in *Advances in Biometrics: Proceedings of Second International Conference on Biometrics (ICB 2007)*. Springer, 2007, pp. 760–769.

[59] S. Draper, A. Khisti, E. Martinian, A. Vetro, and J. Yedidia, "Using distributed source coding to secure fingerprint biometrics," in *Proceedings of the 32nd International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2007, pp. II–129–II–132.

Security for the Smart Grid – Enhancing IEC 62351 to Improve Security in Energy Automation Control

Steffen Fries*, Hans Joachim Hof*, Thierry Dufaure+
Siemens AG
*Corporate Technology; +Energy Automation
Germany
{steffen.fries; hans-joachim.hof; thierry.dufaure}@siemens.com

Maik G. Seewald
Cisco Systems
Hallbergmoos, Germany
maseewal@cisco.com

Abstract— **Information security has gained tremendous importance for energy distribution and energy automation systems over the last years. Security for the smart grid is crucial to ensure reliability and continuous operation of the smart grid. However, the smart grid comes along with new use cases that impose new challenges on existing standards like IEC61850. IEC61850 offers standardized communication services and standardized data models for communication in energy automation, hence it is beneficial for the realization of the smart grid. IEC 61850 is flanked by the standard IEC 62351 that addresses security and specifies technical requirements, which have to be met by vendors. This paper provides an overview about the different aspects of security necessary to build and operate smart grid systems by describing current and new use cases. The focus lies on the current state of the standardization of IEC 62351 and its applicability to the described use cases. Moreover, this work discusses potential enhancements of the standard to address potential shortcomings through changed business and operation models leading to changed trust relations in new use cases like decentralized energy generation and load control. These shortcomings are addressed by describing potential enhancements for part 4 of IEC 62351 allowing multiple parallel distinguishable sessions based on the Manufacturing Message Specification and proper end-to-end authentication as well as authorization.**

Keywords – Smart Grid; Information Security; Cyber Security; Authentication; Authorization; Energy Automation;Smart Home; IEC Standards; NERC-CIP.

I. INTRODUCTION

Power generation and distribution systems are characterized by the existence of two infrastructures in parallel, the electrical grid, carrying the energy, and the information infrastructure used to automate and control the electrical grid. Especially the latter is becoming more and more one of the critical parts of power system operations as it is responsible not only for retrieving information from field equipment but most importantly for sending control commands. A dependable management of these two infrastructures is crucial and strongly relies on the information infrastructure as automation continues to replace manual operations. Hence, the reliability of the power system strongly depends on the reliability of the information infrastructure. Therefore the information infrastructure must be managed to the level of reliability needed to provide the required stability of the power system infrastructure to prevent any type of outage.

The current, rather centralized approach for power generation is evolving to a decentralized power generation involving existing power plants, power plants producing renewable energy (like wind parks) down to households having their own micro power plants (e.g., solar cells). Decentralized energy generation (e.g., solar cells) is believed to become more and more important and common in the future to fight global warming by reducing the CO_2 footprint. Introducing decentralized energy generators into the current energy distribution network poses great challenges for energy automation (EA) in the smart grid scenario, especially secure communication between a control station (e.g., substation) and equipment of users (e.g., decentralized energy generators) must be addressed. Moreover, electro mobility will become more important and needs to be integrated into the current power system landscape. This increases the complexity of power systems even more. In addition, there is also the trend to interconnect the formerly closed and proprietary architectures with office environments and enterprise systems to provide new functionalities and increase cost effectiveness on the move to smarter grid infrastructures. This is accompanied by complete restructuring of the conventional roles on energy market participants.

The classical system architecture of the electric power grid defines distinct roles for energy producers, suppliers and consumers. With the new paradigm of smart grids driving towards sustainability, some of these roles will be redefined. The energy supplier systems have to handle an increasing amount of energy gained from distributed renewable energy sources and independent power production systems in residences. These forms of energy are produced in a much more decentralized way and also have a much more volatile characteristic compared to traditional forms of energy provided by existing power plants, often called bulk generation. At the same time one of the key factors for efficient and economic power generation is a balanced load level on power plants. Smart grid is the approach to address the mismatch between energy generation and consumption. Both aspects directly influence the distribution process of transport and distribution system operators and require the adoption of advanced information and communication technologies (ICT) in these processes.

As the information infrastructure can be described as the backbone of the smart grid and therefore needs appropriate protection to ensure a stable operation of power systems in order to support the required system reliability. Information and cyber security provides the base for protection and resiliency against cyber attacks. This has also be addressed in the comprehensive document set NISTIR 7268 from the Smart Grid Interoperability Panel (cf. [6], [7], and [8]).

The remainder of this paper is organized as follows: Section II provides an overview about energy automation control frameworks focusing on IEC 61850 as one corner stone for the smart grid. Section III discusses security requirements in the context of smart grid. The following Section IV discusses the currently available security in terms of the standard IEC 62351. Based on this, Sections V to VIII discuss potential shortcomings of the standard that become visible through new smart grid use cases. Sections IX and X provide an outlook for potential future work and a conclusion.

II. Energy Automation Control Frameworks

Typical automation systems are built in a hierarchical way. Figure 1 shows typical layers of an automation pyramid. On the lowest level there are sensors and actors like switchgear that are connected to field devices. Serialized field buses as used for a long time are increasingly be replaced by standard communication technology as Ethernet and IP. These field devices are actuated by, e.g., substation controllers, which may be interconnected with other substation controllers using TCP/IP based protocols. On the top are interconnections to supervisory systems the so called control centers, again via TCP/IP.

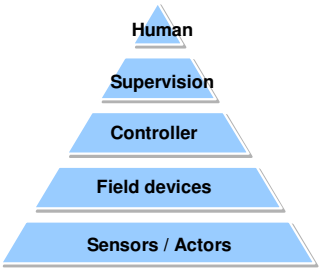


Figure 1. Automation Pyramid

IEC 61850 is a popular standard for communication in the domain of energy automation. It is assumed to be the successor of the currently used standards IEC 60870-4-104 and DNP3 also for the North American region. IEC 61850 enables interoperability between devices used in energy automation, i.e., two IEC 61850 enabled devices of different manufacturers can exchange a set of clearly defined data and the devices can interpret and use these data to achieve the functionality required by the application due to a standardized data model. In particular IEC 61850 enables continuous communication from a control station to decentralized energy generators by using a standardized data format.

IEC 61850 addresses the data exchange on three levels: process level, field level, and station level. It defines the following four important aspects on these levels: Standardized self-describing data, standardized services, standardized networks, and standardized configuration for a complete description of a device. An XML-based system description language – Substation Configuration Language (SCL) – is used to describe a device. Standardized services are used to send standardized data over standardized communication systems. However, IEC 61850 defines abstract communication services that are mapped on existing protocols like TCP/IP, and Ethernet, using the Manufacturing Message Specification (MMS). Moreover, there are also dedicated IEC standards mapping of the IEC 61850 to the target application domain, like IEC 61400-25 providing an adaptation for wind power plants. Here, a mapping to Web Services is targeted and currently under discussion. Security for IEC 61850 is addressed in the related standard IEC 62351 that is described in the following section.

Today, IEC 61850 is mainly used for reporting status and transmitting sampled value information from Intelligent Electronic Devices (IED) to Substation automation controller as well as for command transport from Substation automation controller to IEDs. It also addresses the communication directly between IEDs using the Generic Object Oriented Substation Event (GOOSE) instead of dedicated wires. Necessary tasks comprise also configuration of equipment as well as control of circuit breakers.

The following Figure 2 gives an example of the communication between multiple substations using IEC 61850.

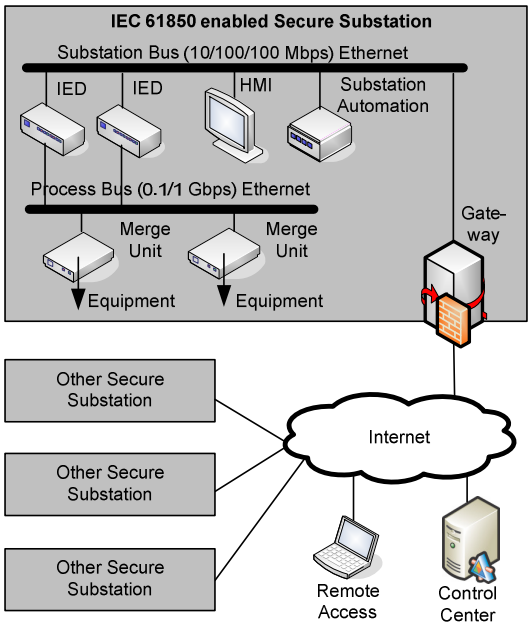


Figure 2. Typical IEC 61850 Scenario

III. SMART GRID SECURITY REQUIREMENTS

Security requirements stem from regulation, technical boundary conditions, and/or direct end-customers. Regulative requirements are given, e.g., by the following regulations.

A. *Regulations and Regulative requirements*

NERC-CIP: North American Electric Reliability Council (NERC) has established the Critical Infrastructure Protection (CIP) Cyber Security Standards CIP–002 through CIP–009 , which are designed to provide a foundation of sound security practices across the bulk power system. These standards are not designed to protect the system from specific and imminent threats. They apply to operators of Bulk Electric Systems (see also [2]). The standards originate in 2006. Last updates have been made in May 2009, but new parts of the standards (CIP 010 and CIP 011) are currently under development.

NERC-CIP provides a consistent framework for security control perimeters and access management with incident reporting and recovery for critical cyber assets and cover functional as well as non-functional requirements. TABLE I provides an overview about the different NERC-CIP parts.

TABLE I. NERC-CIP Overview

CIP	Title / covers
002	**Critical Cyber Asset Identification** Identification and documentation of Critical Cyber Assets using risk-based assessment methodologies
003	**Security Management Controls** Documentation and implementation of Cyber Security Policy reflecting commitment and ability to secure Critical Cyber Assets
004	**Personnel and Training** Maintenance and documentation of security awareness programs to ensure personnel knowledge on proven security practices
005	**Electronic Security Protection** Identification and protection of Electronic Security Perimeters and their access points surrounding Critical Cyber Assets
006	**Physical Security Program** Creation and maintenance of physical security controls, including processes, tools, and procedures to monitor perimeter access
007	**Systems Security Management** Definition and maintenance of methods, procedures, and processes to secure Cyber Assets within the Electronic Security Perimeter to do not adversely affect existing Cyber Security Controls.
008	**Incident Reporting & Response Planning** Development and maintenance of a Cyber Security Incident response plan that addresses classification, response actions and reporting
009	**Recovery Plans for Critical Cyber Assets** Creation and review of recovery plans for Critical Cyber Assets
Draft 010	**Bulk Electrical System Cyber System Categorization** Categorization of BES systems that execute or enable functions essential to reliable operation of the BES into three different classes.
Draft 011	**Bulk Electrical System Cyber System Protection** Mapping of security requirements to BES system categories defined in CIP-010

As already stated, NERC-CIP relates primarily to the operation of critical infrastructure. Nevertheless, this also places requirements on the product vendors to cope with certain security requirements.

BDEW: The "Bundesverband für Energie- und Wasserwirtschaft – BDEW was founded by the federation of four German energy related associations: Bundesverband der deutschen Gas- und Wasserwirtschaft (BGW), Verband der Verbundunternehmen und Regionalen Energieversorger in Deutschland (VRE), Verband der Netzbetreiber (VDN) and Verband der Elektrizitätswirtschaft (VDEW). The BDEW introduced a white paper defining basic security measures and requirements for IT-based control, automation and telecommunication systems, taking into account general technical and operational conditions. It can be seen as a further national approach targeting similar goals as NERC-CIP but less detailed. The white paper addresses requirements for vendors and manufacturers of power system management systems and can be used as an amendment to tender specification.

B. *Supportive actions*

Besides regulative actions, there are also supporting actions, that currently take place, e.g., by investigating in currently available standards and technologies, e.g., by the NIST (National Institute of Standards and Technologies) Smart Grid Interoperability Project (see also [4]). There are two documents, which are mentioned here as they are very compulsory covering a wide range of existing material as well as requirements for further investigation, that have been accomplished by NIST:

- *NIST Framework and Roadmap for Smart Grid Interoperability Standards*, identifying technical standards and specifications, which are also relate to smart grid security (cf. [5]).
- *NISTIR 7628* (cf. [6], [7], and [8]) originates from the Smart Grid Interoperability Panel (Cyber Security WG) and targets the development of a comprehensive set of cyber security requirements building on the NIST SP 1108 (cf. [5]), also stated above. The document consists of three subdocuments targeting strategy (cf. [6]), security architecture (cf. [7]), and requirements, and supportive analyses and references (cf. [8]).

In addition to the NIST activities, the IEC has issued the IEC SG3 report (SMB/4175/R), which encompasses requirements, status and recommendations of standards relevant for the Smart Grid. Security is covered in detail in a separate section of this document. An overall security architecture capturing the complexity of the Smart Grid is requested. Beside this, the following recommendations pertaining open items and necessary enhancements are listed:

- A specification of a dedicated set of security controls (e.g., perimeter security, access control…)
- A defined compartmentalization of Smart Grid applications (domains) based on clear network segmentation and functional zones
- A specification comprising identity establishment (based on trust levels) and identity management
- Security of the legacy components must be addressed by standardization efforts
- The harmonization with the IEC 62443 standard to achieve common industrial security standards
- Finally, it is recommended to review, adapt and enhance existing standards in order to support general and ubiquitous security across wired and wireless connections.

IV. Secure Energy Automation based on IEC62351

Security services to be supported in energy automation comprise the usual suspects:

- **Authentication**: The property that the claimed identity of an entity is correct.
- **Authorization**: The process of giving someone permission to do or have something.
- **Integrity**: The property that information has not been altered in an unauthorized manner.
- **Non-repudiation**: The property that involvement in an action cannot be denied.
- **Confidentiality**: The property that information is not made available or disclosed to unauthorized individuals, entities or processes.

In contrast to office networks, automation networks have different requirements to security services as shown in the following figure.

	Office	EA-Network
Confidentiality (Data)	High	Low – Medium
Integrity (Data)	Medium	High
Availability / Reliability	Medium	High
Non-Repudiation	Medium	High
Component Lifetime	Short - medium	Long

Figure 3. Comparison Office/Automation security

In the context of energy automation, IEC 62351 defines explicit security measures for TCP-based and serial protocols. It applies directly to substation automation deploying IEC 61850 and IEC 60870-x protocols as well as in adjacent communication protocols supporting energy automation, like ICCP (TASE.2) used for inter-control center communication. A clear goal of the standardization of IEC62351 is the assurance of end-to-end security. The standard comprises multiple parts that are in different state of completion.

While part 1 and 2 are more general and comprise the explanation of threat scenarios and the definition of terms, part 3 to 8 are directly related to dedicated protocols like IEC 61850 (IEC 62351 Part 6) and IEC 60870-5-x (IEC 62351 Part 5) and their mappings to lower layer protocols like TCP/IP (IEC 62351 Part 3) and MMS (IEC 62351 Part 4) as well as the mapping of security to the network management (part 7) and role-based access control (part 8). These parts utilize symmetric as well as asymmetric cryptographic functions to secure the payload and the communication link. The remaining part of this section provides an overview about the different parts of IEC 62351 and their current status in standardization.

IEC 62351 applies existing security protocols like Transport Layer Security (TLS, cf. [10]), which has been successfully used in other technical areas and industrial applications, in different parts of the standard. The application of TLS provides for security services like mutual authentication of communication peers and also integrity and confidentiality protection of the communicated data. Thanks to the mutual authentication required by IEC 62351 attacks like Man-in-the-Middle can be successfully countered.

Part 3 of IEC 62351 defines how security services can be provided for TCP/IP based communication. As TLS is based on TCP/IP part 3 specifies cipher suites (the allowed combination of authentication, integrity protection and encryption algorithms) and also states requirements to the certificates to be used in conjunction with TLS. These requirements comprise for instance dedicated certificate context, application of signatures, and the definition of certificate revocation procedures. For the latter, the focus lies mostly on Certificate Revocation Lists (CRL). The application of the Online Certificate Status Protocol (OCSP) is not considered due to limited communication links within the substations. In contrast to office applications, the connections in energy automation are relatively long lasting. This requires the definition of strict key update and CRL update intervals, to restrict the application of cryptographic keys not only for a dedicated number of packets but also for a dedicated time. Another challenge are interoperability requirements between implementations of different vendor's products.

Part 4 of IEC 612351 specifies procedures, protocol enhancements, and algorithms targeting the increase of security messages transmitted over MMS. MMS is an international standard (ISO 9506) dealing with a messaging system for transferring real time process data and supervisory control information either between networked devices or in communication with computer applications. Part 4 defines procedures on transport layer, basing on TLS, as well as on application layer to protect the communicated information. One goal of this paper is to analyze if the defined security is appropriate especially in the context of smart grid applications. This will be discussed in detail in Section VI.

Besides TCP/IP, IEC 62351 Part 5 relates to the specialties of serial communication. Here, additional security measures are defined to especially protect the integrity of the serial connections applying keyed hashes. This part also specifies a separate key management necessary for the security measures.

Part 6 of IEC 62351 describes security for IEC 61850 Peer-to-Peer Profiles. It covers the profiles in IEC 61850 that are not based on TCP/IP for the communication of Generic Object Oriented Substation Events (GOOSE), and Sample Measured Values (SMV) using, e.g., plain Ethernet. Specific for this type of communication is the usage of multicast transfer, where each field device decides based on the message type and sender if it processes the message or not. Security employs digital signatures on message level to protect the integrity of the messages sent, to also cope with multicast connections.

IEC 62351 Part 7 describes security related data objects for end-to-end network and system management (NSM) and also security problem detection. These data objects support the secure control of dedicated parts of the energy automation network. Part 7 can help to implement or extend intrusion detections systems for power system specific objects and devices.

Part 8 of the standard is currently in definition and addresses the integration of role-based access control mechanisms into the whole domain of power systems. This is necessary as in protection systems and in con-

trol centers authorization as well as stringent traceability is required. One usage example is the verification of who has authorized and performed a dedicated switching command. Part 8 supports role-based access control in terms of three profiles. Each of the profiles uses an own type of credential as there are identity certificates with role enhancements, attribute certificates, and software tokens.

The following table provides a short overview about the different IEC 62351 parts and their status in standardization:

TABLE II. IEC 62351 Overview

IEC 62351	Definition of Security Services for	Standardization Status
Part 3	TCP / IP (Profile)	Technical Specification
Part 4	MMS (Profile)	Technical Specification
Part 5	60870-5 and Derivates	Technical Specification
Part 6	IEC 61850	Technical Specification
Part 7	Network Management	Technical Specification
Part 8	Role-based Access Control	Committee Draft
Part 9	Credential Management	New Work Item Proposal

A first glimpse at the current IEC 62351 parts shows that many of the technical security requirements to be applied to energy automation components and systems can be directly derived from the standard. For instance part 3 and 4 explicitly require the usage of TLS. They define cipher suites, which are to be supported as mandatory. These parts also define recommended cipher suites and also deprecate cipher suites, which shall not be applied from IEC 62351 point of view. Note, that the mandatory cipher suites do not collapse with the cipher suites the different TLS versions (1.0 – RFC 2246, 1.1 – RFC 4346, 1.2 – RFC 5246) state as mandatory. IEC 62351 always references TLS v1.0 probably to better address interoperability.

Analyzing the standard more deeply shows that several requirements are provided rather implicit. These requirements relate mostly to the overall key management, which guarantees a smooth operation of the security mechanisms. IEC 62351 uses heavily certificates and associated private keys, e.g., in the context of transport layer protection (using TLS) but also on application layer as in part 6 to secure GOOSE. But to apply this type of credentials, the general handling and life-cycle management including generation, provisioning, revocation, and especially the initial distribution to all participating entities needs to be considered. This is currently underspecified, but has been acknowledged by standardization as important for the general operation but also for the interoperability of different vendor's products. As the standard is extensible a new part, describing credential handling in the context of IEC 62351 services is under development. Moreover, a security architecture, required for building, engineering, and operating power systems is a necessary base to ensure safety and reliability of these systems. Hence further work has been initiated to describe hands-on security architecture guidelines for system engineers and operators to implement, manage and operate power systems securely.

Besides standard enhancements, which have become necessary through findings during the implementation of IEC 62351, new scenarios may also require the further evolvement of already existing or new parts of the standard, to better cope with new use cases. This is the focus of the next section, investigating in new scenarios, which slightly deviate from standard substation automation and thus lead to new security requirements.

V. New Use Cases for IEC 61850 and IEC 62351

Current challenges for the power grid include the integration of fluctuating renewable energy sources, distributed power generation, short interval feedback on users on their energy usage, user indicated demand peaks, and the foreseeable need for the integration of private electronic cars, leading to an even higher energy demand of customers at peak times. A "smarter" grid can meet many of these challenges. With the move to a Smart Grid the importance of IT communication technologies in energy automation rises. With the availability of pervasive IT communication services, a bunch of new use cases become possible that enhance the service to the customer and mitigate the impact of the challenges mentioned above. These new use cases include dynamic pricing, time of use pricing, selling local power into the grid, smart metering, and the like. As IEC 61850 is an introduced standard, the trend is to use this standard to realize these new use cases. While this keeps the effort low to implement new use cases, it may bring new security requirements up that are not addressed by IEC 62351 yet.

A. *Consumer Perspective: Smart Home*

Many use cases center around the Smart Home scenario. Smart Home in combination with the Smart Grid will allow people to understand how their household uses energy, manage energy use better, sell energy produced by local distributed energy generation, and reduce their carbon footprint. IEC 61850 is a natural candidate to use for communication between instances of the Smart Grid and the gateway of a Smart Home.

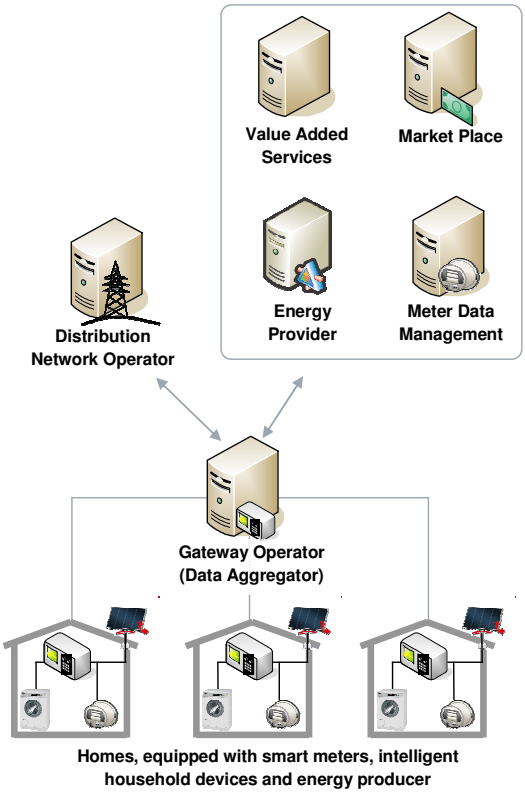


Figure 4. Connection of households to the smart grid

Figure 4 shows a typical system architecture of a smart grid:

- Homes are equipped with smart meters, intelligent household devices, and energy producers.
- Home gateways control the communication between the devices in a home and the Smart Grid and define a security perimeter. The home gateway hides the complexity of the in-house network from the Smart Grid. The home gateway may act as a proxy for the appliances of the home, e.g., on the market place.
- A gateway operator is responsible for administration of the home gateways and provides connectivity for the home gateways.
- The distribution network operator communicated with the home gateways by the means of another instance (in this case, the gateway operator is this instance) that hides the complexity of the home gateway management from the distribution network operator.
- The Meter Data Management manages the metering data received from the smart meters. The Meter Data Management processes the metering data for the various energy providers and provides them with a summary for accounting.
- At the energy market, consumers (resp. their home gateways) buy energy, and energy generators sell energy; hence the market offers a demand regulated price. An energy market alleviates the integration of distributed energy generators (e.g., solar cells).
- The smart grid communication infrastructure and the energy market are the enabler for other value added services.

Having a communication and IT infrastructure like this at hand, the following use cases are possible in a Smart Home scenario:

1) *Energy-aware home appliances*

Nowadays, the price of energy for private consumers is mostly constant. From the perspective of a utility it would be beneficial to have dynamic pricing to influence the energy usage of customers. On the customer side, new intelligent, energy-aware home appliances can optimize the costs for energy usage by starting and stopping energy extensive tasks (e.g., cloth or dish washing) at appropriate times (e.g., start when energy is cheap). This requires that the current price of energy is known and there is some way to determine the price of energy for the duration of an operation (e.g., washing a load of wash). One way to implement such a system is an energy market, where energy-aware home appliances buy a certain amount of energy before they start an operation. Especially charging a private electrical car during the night is an extremely flexible operation that requires much energy but has a large time window for execution, hence benefits from a good deal.

To implement this use case with the architecture presented above, the home gateway trades energy at the energy market. Accounting for any contract on the energy market includes the energy provider as well as the meter data management.

2) *Distributed power generation*

If energy is produced in a home, e.g., by solar cells,

the energy is traded on an energy market to achieve the best possible price. Especially if the energy market is on a large scale, selling the energy may be more attractive than a fixed pricing.

To implement this use case with the architecture presented above, the home gateway trades energy at the energy market. Accounting for contracts includes the distribution network provider as well as the smart meter management.

3) Energy Management and User Awareness

An application with integrated user interface in the home is used for communication with the utility, e.g., to get a diagram of current energy usage, to get current energy pricing, to get the personal energy usage history, to get energy saving tips and the like. The user interface may also be used to receive energy outage forecasts, for troubleshooting, or to dynamically select a desired energy mix.

Even energy-aware home appliances may offer a user interface that states the current price for one operation execution. E.g., a coffee machine may state the price per coffee pot.

To implement this use case with the architecture presented above, the home gateway informs appliances about current energy prices, which it either gets at the energy market or directly from the energy provider (price signals as special incentive for special behavior).

B. Utility Perspective

Other use cases are focused on keeping the distribution network stable and keeping costs for utilities low (e.g., because it is not necessary to buy additional energy at short notice). As IEC 61850 is already widespread in use in the distribution network, it is a natural candidate for the following use cases:

1) Reactive shutoff of home appliances

A utility has the ability to shut down certain home appliances in the household of users on short notice to react on certain situations in the network (e.g., if too many consumers are active). Such switch-off commands can be based on special contracts between user and utility operator.

To implement this use case with the architecture presented above, the utility must have a list of home appliances that can be shut off as well as the communication addresses of the associated home gateways. In the architecture above, home gateways may be addressed by the gateway operator that also ensures the connectivity of the home gateways. The utility sends a shutoff message via the gateway operator to a set of home gateways. Sending this shutoff message to many households must be finished in a short time to allow fast reactions. The shutoff message must be protected to avoid being misused by attackers. The home gateway takes the appropriate actions to meet the request of the utility, especially, it communicates with proper appliances to be shut off.

2) Shutoff of power generator

The utility may not only turn off certain home appliances, it may also instruct distributed power generators not to feed energy to the distribution network to fight situations when there is a low demand for energy. The signaling process is the same as in the last use case.

3) Demand Response

Another use case from a utility prospect is demand response: A utility can send price signals (either a rather high price if energy demand is too high or a low price if the energy demand is too low) to influence energy usage of intelligent home appliances without using the energy market. Price signals are especially interesting for the loading of electric cars. Price signals can be sent for future time periods or as real time pricing information. The utility sends price signals via the gateway operator that knows to address the home gateways. The home gateways distribute the pricing information in the home to the appropriate home appliances.

4) Asset Management

Yet another use case from the utility perspective is asset management. Given a rising number of equipment for decentralized energy generation in the households of the users, managing the network gets more complex. An automated asset management helps to reduce costs and gives a good view on the state of the distribution network. IEC 61850 includes self-describing configurations of device and all kind of tracking data; hence it is a natural candidate for the following use cases:

- Utilities collect data about the state of the network and about the equipment in a user's home.
- Utility gathers circuit and/or transformer load profiles, makes decisions on asset replacement based on a range of inputs including comprehensive off line and on line condition data and analysis
- Utility performs localized load reduction to relieve circuit and/or transformer overloads

– Utility system operator determines level of severity for an impending asset failure and takes corrective action

C. New Requirements

One requirement arising from these new use cases is scalability. Security solutions for the Smart Grid must scale with millions of devices - Germany for example has more than 39 million households and each household may have more than one device. Multiple levels of hierarchy from a control station to a device in a household are a common solution to address scalability. This includes communication other than the point to point communication used today.

As shown in Figure 4, in smart grid scenario's new roles and/or components may be introduced in terms of a home energy gateway operator. This gateway operator is in charge of concentrating the communication from the home energy gateways up to the control center as well as providing an easy way to the control center to reach a high number of energy gateways at once. Moreover, a gateway operator may also offer additional services like remote management of the home energy gateways, e.g., to provide enhanced functionality or path and updates for installed software. This new component changes the trust assumptions for the substation communication as it may be seen new intermediate component, which belongs to a different security domain. This component most likely terminates the transport connection between a control center and the home energy gateway, which is used synonym here for a field device.

Today's security solutions assume trusted intermediate nodes if one application connection is realized over multiple transport connections. This assumption may not hold in the future and new security concepts may only assume intermediate nodes that forward traffic but may or may not be trusted.

The following section targets the analysis of applying IEC 62351 in the context of the smart grid scenario just described to discuss, if the standardized security provides sufficient counter measures.

VI. MISSING PIECES IN IEC62351

As stated in Section II above, part 4 of IEC 62351 specifies procedures, protocol enhancements, and algorithms targeting the increase of security of applications utilizing the MMS. MMS is an international standard (ISO 9506) dealing with a messaging system for transferring real time process data and supervisory control information either between networked devices or in communication with computer applications. Within IEC 61850 there exists a mapping to MMS to transport commands and data between the different energy automation components. Thus IEC 61850 can directly leverage the security enhancements defined in part 4 of IEC 62351.

The security, as defined in IEC 62351 part 4, is described by two profiles targeting transport security as T-Profile on one hand and application security as A-Profile. The T-Profile describes the protection of information, which is exchanged over TCP/IP using TLS. This is mainly being done by referring part 3 for TLS application and the definition of additional mandatory cipher suites. The A-Profile defines security services on application layer, targeting mainly authentication. Note that the authentication in the A-Profile is performed only during connection establishment on application layer using the MMS initiate command. Moreover this authentication is defined in a way that it does not provide application layer message integrity. Furthermore the authentication phase is not used to form a session. A session in this context cryptographically binds the authentication performed during the connection setup with subsequent messages exchanged between the communicating peers. Thus, in the current stage of the standard, messages on application layer are not protected regarding their integrity. To achieve integrity protection, the application of the T-Profile is being referred.

Combining A-Profile and T-Profile provides a connection allowing for authentication, integrity protection and confidentiality on transport level and authentication on application level. This approach works fine in scenarios, where the transport connection spans the same entities as the application connections as shown in Figure 5.

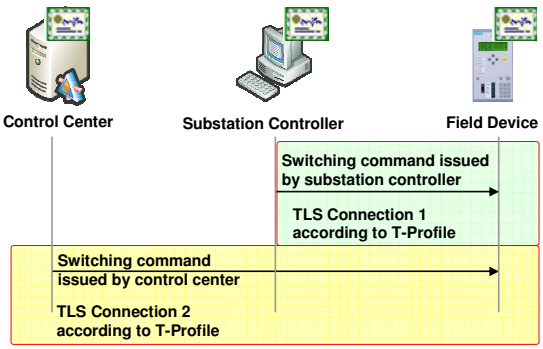


Figure 5. Direct switching action

While this approach may be sufficient for many energy automation scenarios, it may not cope with new use cases, for instance the ones described in Section V.

As soon as there is a difference in transport connection hops and application connection hops, security problems may arise. An example may be a scenario in which a proxy is used, e.g., to combine different connections or to multicast a single command to several other connections as described in Figure 4 by the gateway operator. From the standard energy automation architecture – Control Center, Substation Controller, Field Device – this gateway operator resembles the substation controller and operates as a communication proxy as shown in Figure 6. Therefore, the T-Profile is terminated by the substation controller, while the application connection may be established end-to-end, directly with the actual entity to be reached. Since IEC 62351 part 4 does not provide application level integrity, no end-to-end application level security is provided.

Such a scenario can be described as multi-hop connection from a transport level view and would require that the proxy is a trusted intermediate host, which cannot be guaranteed in many scenarios. For example in one of the new use cases addressed in the last section, a utility may use a number of proxy that multicasts a single "switch off" command issued by the control station to multiple households. This approach allows multiple hierarchy level for issuing the "switch off" commands to achieve scalability and fast reaction.

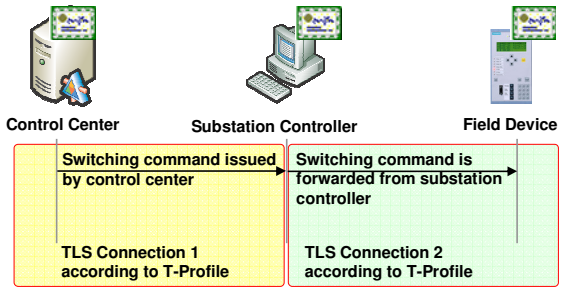


Figure 6. Proxied switching action

To provide also end-to-end integrity in multi-hop use cases with intermediate nodes additional measures have to be defined. Ideally, these will enhance the standard IEC 62351 to foster both, security and interoperability.

The approach to find appropriate security enhancements taken here involves the investigation into existing protocols, which already provide a secure session concept on application layer. The following section analyzes different approaches to enhance part 4 based on existing security measures.

VII. CANDIDATES FOR ADAPTATION

This section discusses three potential candidates, which are already defined and widely used in communication technology and their suitability for IEC 62351 part 4 to better cope with multi-hop scenarios. As stated in the previous section, the additional security requirements to be met comprise peer authentication and message integrity on application layer between end-to-end communicating peers. The three candidates are:

– HTTP Digest Authentication as typically used in web based communication

– H.235 based security as used to protect multimedia communication

– XML security as applied in web service frameworks

The goal is the enhancement of MMS communication to allow cryptographically based sessions to provide end-to-end security on application layer. Moreover, being able to associate MMS commands with a dedicated session, also allows running multiple parallel distinguishable sessions over the same T-Profile protected link(s).

A. *Candidate 1 HTTP Digest Authentication*

RFC2617 (cf. [9]) describes authentication options in the context of HTTP (Hypertext Transport Protocol), which is used in many web-based applications. While basic authentication is deprecated because of its worst security, digest authentication is being widely used. In digest authentication a shared secret needs to be available on both ends of the communication, which is used to calculate an MD5 checksum over either a certain part of the message or the complete message as part of a challenge response mechanism to provide integrity protection. Typically each HTTP request can be challenged to authenticate the requestor. In the worst case this would mean that each communication action is doubled. The general approach is depicted in Figure 7.

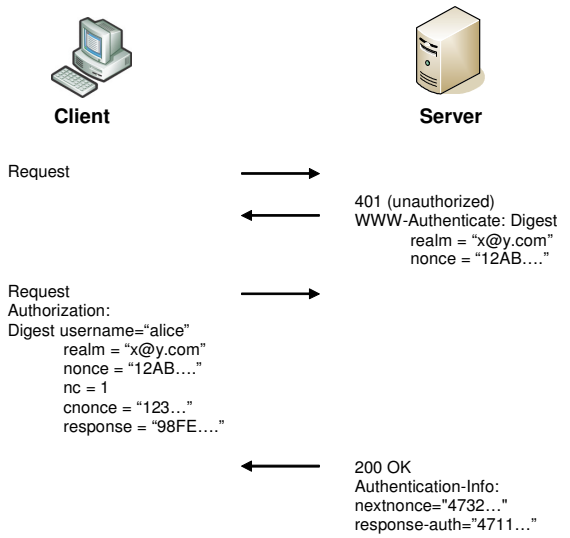


Figure 7. HTTP Digest Authentication

To avoid the doubling of all message exchanges the challenge for the next exchange can be transmitted as part of the response message to the initial request as optimization f this method. The next nonce mechanism in combination with the initial application of username and password can be used to form a (weak) cryptographic session.

B. Candidate 2 H.235 based security

H.323 is an umbrella recommendation defined by the ITU-T (International Telecommunication Union) to address call control, multimedia management, and bandwidth management in telecommunication environments. H.235 is also an ITU-T based standard describing security functions for the multimedia communication standard H.323. H.235 features in summary nine different profiles, were only some of them are interesting to be discussed in the context of leveraging them for the securing of MMS:

- **H.235.1** provides signaling integrity and authentication using mutually shared secrets and keyed hashes, based on HMAC-SHA1-96. This profile is widely implemented in available H.323 solutions.
- **H.235.2** provides signaling integrity and authentication using digital signatures on every message in gatekeeper-routed scenarios. Since signature generation and verification is costly in terms of performance, this profile may not gain momentum and is stated here rather for completeness.
- **H.235.3** is a hybrid approach using both, H.235.1 and H.235.2. During the first handshake a shared secret establishment is performed, protected by digital signatures. Afterwards keyed hashes are used for message integrity protection, based on the established shared secret.

The syntax of the H.323 messages is depicted in Figure 8. As it can be seen, security is provided based on an included crypto token in the message, which transports all necessary data to integrity protect the message.

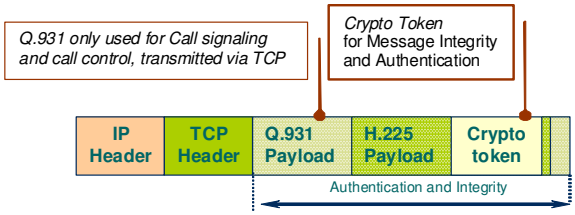


Figure 8. H.235 protected message

As H.235.3 allows for a hybrid security approach, utilizing asymmetric and symmetric cryptography, the crypto token is defined to serve for both approaches at once and carries all necessary information for both phases.

C. Candidate 3 XML Security

The Extensible Markup Language (XML) is a simple, very flexible text format, which is defined by the W3C (World Wide Web Consortium). It specifies a set of rules for encoding documents in machine-readable form and is meanwhile used in a variety of applications and builds the base for message structures in several protocols and language derivations.

The W3C also provides recommendations for security of XML data. XML security comes in two flavors, XML Encryption and XML Signature. Both can be used on XML encoded data in so-called XML elements and provide privacy and integrity protection. XML encryption allows the encryption of any type of data with symmetric and asymmetric methods. The key to be used can be selected by key names. XML signature on the other side applies asymmetric methods to achieve integrity protection and non-repudiation and can be included in the XML document directly or provided in a detached fashion (see also [18]).

VIII. Proposed Enhancements of IEC62351

Based on the discussion of candidates in the previous section and the fact that integrity protection is the first protection goal in energy automation networks, the approach of candidate 1 and 2 and their application to

MMS is discussed here further, as they allow the integrity protection of application layer messages based on an cryptographic authenticated and integrity protected session. The application of a hybrid approach as in candidate 2 in this context, using asymmetric key material for the authentication and protection of a session key establishment and symmetric key material for the remaining session provides for a high flexibility while keeping the load on the system low during the application of the symmetric key. This cannot be achieved with candidate 1.

Candidate 3 is not discussed further here as directly it maps to web services instead of MMS. IEC 61400-25 (for wind power plants) describes a mapping of IEC 61850 services to web services. Moreover, other approaches like OPC-UA (Object Linking and Embedding for Process Control – Unified Architecture) also apply web service technology and may also be used in this context. As for web services own security measures are defined (e.g., XML security), these security measures may be applied straight forward. Nevertheless, these possibilities should be kept in mind, to provide an adequate security level for MMS, operating at the same level as web-services. This is especially important for the protocol interworking when different transport mappings are used.

Again, the goal is the enhancement of MMS-based communication to allow multiple parallel distinguishable integrity protected sessions started with the MMS Initiate command and proper authentication (and authorization).

Providing this security session approach can generally be done in different ways:

1. Enhancement of commands transported via MMS with security tokens to allow authentication and authorization to be bound to the messages directly. This approach would be independent of MMS security and thus may be applied over other transports as well.
2. Enhancement of MMS itself to allow security services on the layer transporting IEC 61850 commands. This approach requires fewer changes in the current message structure and better interoperates with other approaches, like security options for web services.

The enhancement of the MMS messages itself requires changes in IEC62351 Part 4 for security of MMS communication as currently only the MMS initiate command has the appropriate ASN.1 structures to transport the security information. It also requires changes in the IEC 61850 standard to provide the necessary integrity field carrying the security parameters as a base for the introduction of a cryptographic session concept.

Therefore, the current approach of MMS must also be enhanced to provide not only authentication, but also integrity protection. This means the current description of the signature calculation in IEC62351 Part 4 needs to be revised.

The following discussion relates to candidate 1 and 2 explained in the previous section:

The basic idea for both approaches, the enhancements of the syntax of the commands send via MMS (case 1 above) or of the MMS message syntax (case 2 above), is the enhancement of the datagram with a substructure to transport all necessary security information. This change may be done as Figure 9 suggests, based on the investigation into the realization of candidate 2 in the previous section.

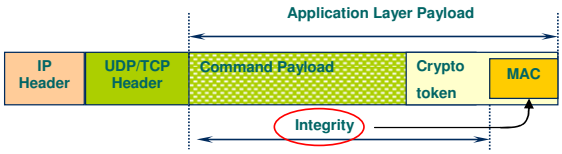


Figure 9. Message protection using a crypto token

The application of a *crypto token* provides a dedicated security container to transport message authentication codes and additional information, e.g., necessary to setup a session key.

An alternative addressing only message integrity on application layer without enabling the transport of key establishment values for the integrity protection is depicted in Figure 10. This approach would be suitable, when focusing on candidate 1.

Figure 10. Message integrity protection

For the following discussion, the approach using a *crypto token*, as depicted in Figure 9 is favored as it offers most flexibility. The command payload may be seen either on MMS level (Layer 6) or on application level (Layer 7). In any case, the *crypto token* to be included in the payload carries at least (necessary pa-

rameter should be discussed, depending on the solution approach; the following list may not be complete):

- tokenOID Object identifier
- certificate certificate information
- timestamp Timestamp
- sequence Sequence number
- random nonce value
- dhkey Diffie Hellman set (to negotiate a session key)
- receiverID Receiver Identifier
- sendersID sender Identifier
- hashed message authentication code based on keyed hash (HMAC)
- signed message authentication code based on signatures

The inclusion of the *crypto token* in the messages enables the following functionality:

1. Authentication of connected to and connecting peer during first message exchange, here during the MMS Initiate. Based on the chosen credentials, this may be done using either symmetric or asymmetric long term keys (hashed or signed).
2. Negotiation of a session key during the first handshake to be used for all subsequent messages in this session. This may be done by using for instance the Diffie Hellman Key Agreement, were both, the client and the server provide to the session key. The session enables the distinction of messages sources in terms of applications or users.
3. Integrity protection of messages on application layer. In scenarios, were multiple hosts are traversed this approach does not require to trust an intermediate hosts to not alter messages contents. The intermediate hosts needs only to be trusted to deliver the message.
4. Replay protection through the use of timestamps and sequence numbers or nonce's alternatively.

A potential call flow between a control center and a field device via a substation controller using the described approach of candidate 2 using the MMS layer is shown in Figure 11. This figure also merges the existing energy automation systems with roles and systems of smart grid scenarios with residential integration as shown in Figure 4.

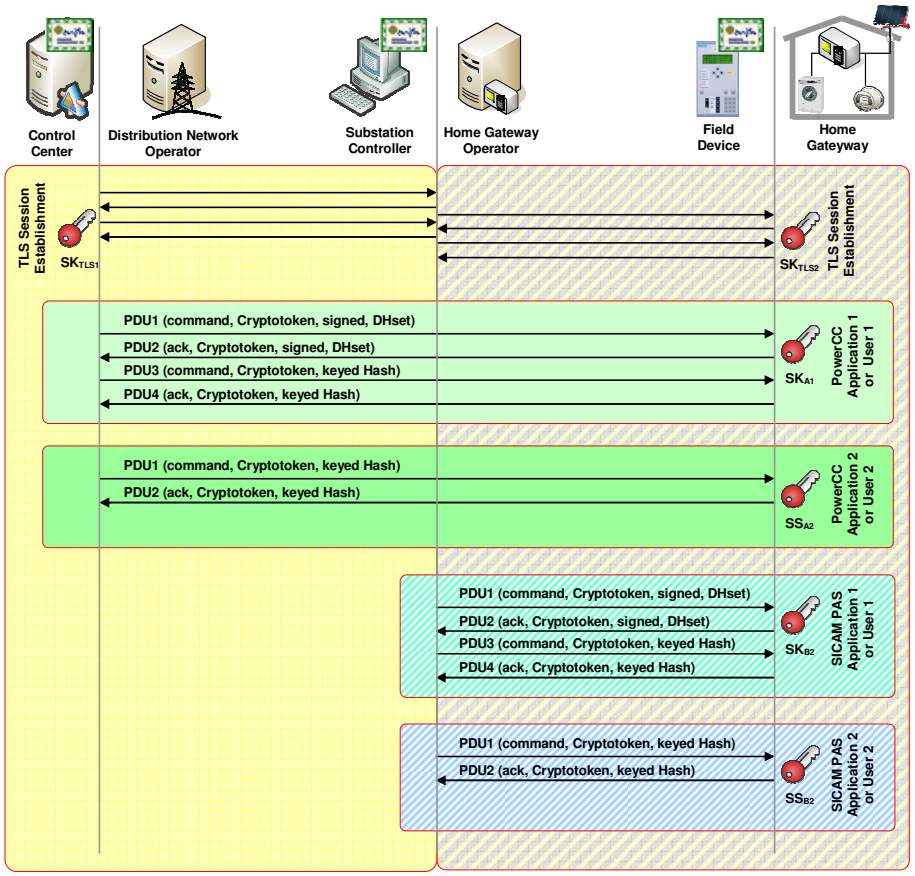


Figure 11. Security enhanced call flow

The following description explains this call flow:

- At first a TLS connection is established on both hops. Here, TLS negotiates session keys on transport level on both hops: SK_{TLS1} and SK_{TLS2}.
- Afterwards an application/user A1 on the control center issues a command to the field device. As this is the first command for this application/user, the command is authenticated using the long term credential (e.g., digitally signed). The acknowledgement in turn is secured using the long term credential of the field device. During the handshake a session key may be established SK_{A1} using a Diffie Hellman key agreement. This session key may then be used to secure all subsequent traffic between A1 and the field device. The command is send via the TLS protected hops via the substation controller to the field device.
- A second application/user A2 on the control center issues a further command to the field device. As both communication parties possess a shared secret SS_{A2}, it is used to secure the message exchange applying a keyed hash (e.g., HMAC-SHA1). The command is send via the same TLS protected hops via the substation controller to the field device.
- Then an application/user B1 on the substation controller issues a command to the field device. As this is the first command for this application/user, the command is authenticated using the long term credential (e.g., digitally signed). The acknowledgement in turn is secured using the long term credential of the field device. During the handshake a session key may be established SK_{B1}. This session key may then be used to secure all subsequent traffic between B1 and the field device. The command is send via the TLS protected hop to the field device.
- A second application/user B2 on the substation controller issues a further command to the field device. As both communication parties possess a shared secret SS_{B2}, it is used to secure the message exchange applying a keyed hash (e.g., HMAC-SHA1). The command is send via the same TLS protected hops via the substation controller to the field device.

The advantage of this approach is that single TLS connections can be used on the hops to secure the transport between all involved peers, while multiple applications or users may use these TLS connections to transport specific commands to the field devices. Moreover, due to the session concept, the long term credentials need only to be used during the first handshake, while all other communication can rely on the negotiated session keys. If digital signatures are performed during the first handshake, performance can be saved on all further messages of this application connection, as the keyed hash operation is less consuming compared to a signature generation or verification. The approach as shown in Figure 11 is suitable for both, MMS or direct command integration.

IX. FUTURE WORK

As already stated in chapter VI, Web Services are gaining more momentum. They have already been addressed as part of the wind power craft related standard IEC 61400-25 and it is expected that there will be a mapping for IEC 61850 in the near future. Web services are also one building block in the OPC-UA framework initially mentioned were security functions already being considered on transport and application layer.

Web services enable the application of Web security mechanisms like XML Security to provide encryption and integrity protection. Moreover authorization can also be addressed utilizing the Security Assertion Markup Language (SAML). SAML allows the definition of secured tokens, to be issued by a trusted component. Currently, security is also not being addressed in the wind power standard. Nevertheless, as web service security is already defined (by the W3C), the standard only needs to be enhanced with a mapping to the available web security, without the necessity to defined own security mechanisms.

To ensure security interworking between installations utilizing different mappings of IEC 61850 like MMS or Web Service secure services transition functions need to be defined. Therefore, from the interworking perspective, the integration of security enhancements in MMS may provide a better base for secure interworking as it operates on the same level as web services and already provides an end-to-end application layer connection.

X. CONCLUSION

This paper provides an overview of smart grid environment focusing especially on the security of dedicated new scenarios, which become more likely through the integration of renewable energy sources not only on substation level, but also on end-user level. Additional security requirements will be the result of these new use cases. The energy automation security standard IEC 62351, which is used to secure communi-

cation according to the standards IEC 61850 and IEC 60870-x and to provide End-to-End Security plays a major role here. Because of the manifold Smart Grid activities and the standardization efforts driven by NIST, new parts of IEC 62351 can be expected soon. Motivated by the analysis of new use cases for Smart Grids, some shortcomings of IEC 62351 are presented. Especially, IEC 62351 can currently not offer application layer end-to-end security if multiple transport layer connections are used. Such multi-hop connections are important for new use cases. Currently, often a trusted intermediate is assumed for application layer end-to-end security. This assumption may be a weakness in the overall system design depending on the use case and may not hold in the future.

An extension of IEC 62351 is proposed to overcome the identified weaknesses by introducing security sessions for MMS connections in IEC 62351. The extension enables cryptographic sessions on application layer providing application layer end-to-end security for new use cases in Smart Grid scenarios.

REFERENCES

[1] Fries, S.; Hof, H-J; Seewald, M.: The Fifth International Conference on Internet and Web Applications and Services – ICIW 2010: "Enhancing IEC 62351 to Improve Security for Energy Automation in Smart Grid Environments", May 2010, ISBN 978-0-7695-4022-1

[2] NERC, North American Reliability Corporation, last access February 2011: http://www.nerc.com/page.php?cid=2|20

[3] BDEW – Bundesverband der Energie- und Wasserwirtschaft, Datensicherheit, last access January 2011: http://www.bdew.de/bdew.nsf/id/DE_Datensicherheit

[4] NIST, National Institute of Standards and Technlogies, Smart Grid Interoperability Project, last access January 2011 http://www.nist.gov/smartgrid/

[5] NIST Framework and Roadmap for Smart Grid Interoperability Standards, Version 1.0, last access January 2011, http://www.nist.gov/public_affairs/releases/upload/smartgrid_interoperability_final.pdf

[6] NIST IR 7628 Guidelines for Smart Grid Cyber Security, Vol. 1 Smart Grid Cyber Security Strategy, August 2010, last access January 2011: http://csrc.nist.gov/publications/nistir/ir7628/nistir-7628_vol1.pdf

[7] NIST IR 7628 Guidelines for Smart Grid Cyber Security, Vol. 2 Security Architecture and Security Requirements, August 2010, last access January 2011: http://csrc.nist.gov/publications/nistir/ir7628/nistir-7628_vol2.pdf

[8] NIST IR 7628 Guidelines for Smart Grid Cyber Security, Vol. 3 Supportive Analyses and References, August 2010, last access January 2011: http://csrc.nist.gov/publications/nistir/ir7628/nistir-7628_vol3.pdf

[9] RFC2617: HTTP Authentication: Basic and Digest Access Authentication, J. Franks, P. Hallam-Baker, J. Hostetler, S. Lawrence, P. Leach, A. Luotonen, L. Stewart, June 1999

[10] RFC 5246: The Transport Layer Security (TLS) Protocol, Version 1.2, T. Dierks, E Rescorla, August 2008

[11] ISO-IEC 61850, Part 1: Introduction and Overview, May 2003

[12] ISO-IEC 61850, Part 8-1: Specific Communication Service Mapping (SCSM) – Mappings to MMS (ISO 9506-1 and ISO 9506-2) and to ISO/IEC 8802-3, May 2004

[13] ISO-IEC IEC 61400, Part 25-4: Communications for monitoring and control of wind power plants – Mapping to communication profile, August 2008

[14] ISO-IEC 62351, Part 4: Communication Network and System Security – Profiles Including MMS, October 2006

[15] ISO-IEC 62351, Part 5: Security for IEC 60870 and Derivatives, February 2007

[16] ISO-IEC 62351, Part 6: Security for IEC 61850, October 2006

[17] H.235.0: Security framework for H-series, ITU-T, 2005

[18] XML Signature Syntax and Processing (second Edition), W3C Recommendation, 10.June 2008, last access January 2011: http://www.w3.org/TR/xmldsig-core/

An Evaluation of BOF4WSS and the Security Negotiations Model and Tool used to Support it

Jason R. C. Nurse and Jane E. Sinclair
University of Warwick, Coventry, CV4 7AL, UK
{*jnurse, jane.sinclair*}*@dcs.warwick.ac.uk*

***Abstract*—As online collaboration between businesses increases, securing these interactions becomes of utmost importance. Not only must entities protect themselves and their electronic collaborations, but they must also ensure compliance to a plethora of security-related laws and industry standards. Our research has focused in detail on the cross-enterprise security problems faced by collaborating businesses. Apart from our most recent work which investigates a novel model and tool to support e-businesses' security negotiations, we previously defined a comprehensive development methodology to aid companies in creating secure and trusted interactions. This paper aims to advance those proposals by presenting and discussing a key stage of their evaluation. This stage uses interviews with industry-based security professionals from the field, to gather critical, objective feedback on the use and suitability of the proposals in fulfilling their aims.**

***Keywords*-Business-oriented framework, e-business collaborations, security negotiations, security ontology, XML security language, interview evaluation**

I. Introduction

Inter-organizational e-business, endorsed by a wide suite of enabling technologies (e.g., Web services, ebXML, RosettaNet), is now one of the most promising and lucrative business paradigms. To sustain these online interactions, security researchers and professionals have investigated numerous technologies, processes and best practices. Apart from our most recent research in [1], in previous other work we have also contributed to this area by defining the Business-Oriented Framework for enhancing Web Services Security for e-business (BOF4WSS) [2], [3]. BOF4WSS' uniqueness stems from its emphasis on a detailed cross-enterprise development methodology, to aid collaborating e-businesses in jointly creating secure and trusted interactions. This particularly refers to the creation of a multilayered security solution, which encompasses technologies, processes, policies and strategies, and spans the interacting companies.

Further to the comprehensive guidance supplied by BOF4WSS, our research has explored the provision of a range of useful support systems. These would assist in the framework's application to business scenarios, and seek to streamline various essential, but often arduous or problematic development tasks. One such support model and resulting system, which we recently developed can be seen in [1]; formally, this paper extends that work. That proposal specifically targeted the difficulties incurred during companies' negotiations on security actions and requirements; a prerequisite activity before the joint systems are developed. Here, a *security action* is defined as any high-level way in which a company handles a risk it faces (e.g., 'the risk of ensuring the security of a server is to be outsourced'), whereas a *security requirement* is a high-to-medium level desire, expressed to mitigate a risk (e.g., 'the integrity of personal data must be maintained'). Security actions thus encompassing security requirements.

The problem area highlighted above and discussed in subsequent sections, relates to the organizational, practical and physical hardships incurred when transitioning from the individually completed Requirements Elicitation stage, to the subsequent Negotiations stage in BOF4WSS. In this latter stage is where interacting companies meet to present, negotiate and reconcile their security actions and requirements. Attempting to address these hardships, the Solution Model and resulting tool for security negotiations support in [1] were created. These proposals specially aimed at streamlining various negotiations tasks and significantly easing framework phase transition for parties. Initial evaluation results in [1] and to a larger extent in [4] have provided a good start in demonstrating Model and tool compatibility with existing security approaches used in businesses.

Having defined the framework and outlined a key support tool in previous works, this paper aims to report on the findings from one of the more substantial, initial evaluation stages. This stage used in-depth interviews with industry-based security professionals from the field, to gather critical, objective feedback on the use and suitability of the proposals in fulfilling their aims. Another prime goal of this evaluation was to gain further insight into industry and business scenario realities before planning and conducting the final evaluation of BOF4WSS and the supporting tools. This final evaluation would constitute a thorough case study analysis.

This paper is organized as follows. Section II recaps BOF4WSS inclusive of its aims and the goals of its phases. This review was seen necessary in the interest of completeness considering the detailed analysis of the framework in the forthcoming evaluation. Work in [2], [3] form the main references for the framework's review. Next, Section III assesses the difficulties incurred in cross-enterprise security

negotiations, and discusses the Model and tool proposed to tackle them. With the main proposals outlined, Section IV reports on the interview-based evaluation of both the framework and the Model and tool. The feedback gathered will be an important finding regarding the use and suitability of the proposals. Conclusions and future work are presented in Section V.

II. The Framework

BOF4WSS [2], [3] is an approach for cross-enterprise security and trust within e-businesses that employ Web services (WS) technology. The prime novelty of this framework is found in its emphasis on providing an expanded formalization of a development methodology that focuses on security and trust. This methodology also accommodates multiple autonomous businesses working together. There are two main shortcomings of existing approaches targeted by the framework. These stem from: (i) an overly reliant emphasis on technology, alluding to standards and systems as the complete solution to WS security in e-business; and (ii) an overly isolated security stance, focusing on the process *one* company should follow to secure itself internally, therefore ignoring the cross-enterprise security issue (discussed in Hartman et al. [5]) introduced by WS use.

To address these outstanding issues, BOF4WSS aims at three aspects. First, to consider the full nature of WS and its security implications within e-business. Second, appreciating that security, irrespective of the context, is a multilayered phenomenon encompassing aspects such as practices, processes and methodologies, in addition to technologies. And finally, to promote the use of a collaborative approach to provide enhanced levels of security and trust across partnering companies.

As seen in [3] and depicted in brief below, the framework and its phases give detailed guidance on what should occur and how, and its pertinence in attaining desired levels of holistic security for these cross-enterprise interactions. This will involve defining the expected inputs to stages, along with their required outputs/outcomes, but especially the recommended low-level goals, activities, and steps within those stages that can help achieve the outcomes. Where suitable, this guidance aims to reuse existing methods and practices—both from industry and academia—thus concentrating on the compilation of these into a coherent, well-defined process.

With the framework's background discussed, Figure 1 displays a pictorial representation of its nine phases. These are then described.

The first phase is **Requirements Elicitation** and within it each business works largely by itself. The tasks conducted include analyzing internal business objectives, constraints, relevant laws, security polices and so on, to determine their high-level needs for the foreseen WS business scenario. Existing methods such as those proposed by Demirörs [6] are used to aid in this task. This technique (that is, [6]) focuses

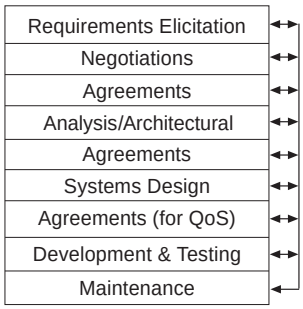


Figure 1. BOF4WSS Overview

on the definition and analysis of business process models to elicit requirements. This type of approach is preferred mainly due to its innate emphasis on business processes—the culmination of the expected service interactions.

In the **Negotiations phase** next, teams consisting of project managers, business and systems analysts, domain experts, and IT security professionals from the companies meet, bringing together their requirements from the previous phase for discussions. The purpose is to use the stage inputs as a basis to chart an **agreed** path forward especially considering the varying expectations each company is likely to have towards security. Expectations (and requirements) could vary with regards to whether a process (or set of service interactions) needs to be secured, to what level is it to be secured, how will security be applied, and so on. Work in [7] clearly highlights that in forming these partnerships of companies, this integration task is formidable. Nonetheless, this is a pivotal step in engaging in interactions.

The **Agreements phase** which follows, uses the completed negotiations to clearly define agreements thus far. The first task suggested by the framework is a legal contract to cement the understanding of the requirements between companies. This legal document is followed by a novel construct called the Interaction Security Strategy (ISS). The ISS as opposed to the contract, is a less rigid management structure that defines high-level, cross-enterprise security directives to guide the interactions. This would form the basis for all the scenario's security decisions instead of individual company's policies or requirements. Another prime goal of the strategy is fostering trust amongst business partners through predictability and transparency in security approaches, by outlining a structure that all entities agreed to adopt and follow. This trust aim is discussed in more detail in [3].

Within the **Analysis/Architectural phase**, the aim is to enable businesses to draw upon the previously agreed requirements and jointly define conceptual business process models for the expected interactions. The directives (policies, best practices, and so on) from the ISS are also then applied to create secure process models. This stage's expected output is a blueprint for the high-to-medium level

process flow and respective security architecture.

Following formal process definition, BOF4WSS advises the use of another **Agreements phase**. This time the goal is towards a more thorough legal contract reflecting detailed requirements and expectations of the companies involved. At this point, contracts are used primarily as a safety net, and should leave the role of governing day-to-day interactions to the ISS.

The aim of the **Design phase** is aiding businesses in defining a logical, low-level systems view of exactly how the conceptual model from the Architectural phase will be achieved. Examples of objectives that constitute this aim are the identification of relevant WS standards, trade-off analysis of their use, and the actual standards application where appropriate. In addition to standards agreement, harmonizing process and data semantics is also an issue worthy of consideration when discussing inter-company interactions as stressed in Papazoglou [8]. A semantics framework including shared vocabularies are therefore to be specified in this framework phase. On the completion of these tasks, the stage is complete. A specification document is therefore output that is appropriate for systems and software developers to implement.

With the low-level processes and functional services specified, the subsequent phase focuses on the **Agreements** at the lower, quality-of-service (QoS) level. The goal is to specify the mutual understanding of the priorities, responsibilities, and guarantees expected by each company regarding the actual Web services. QoS elements typically emphasized encompass performance requirements (e.g., average response time of 30 milliseconds),service availability needs (e.g., uptime of 99.96%), and so on. Apart from formal natural language statements which form what is commonly known as a Service-Level Agreement (SLA), this specification is done using relevant policy and service agreements WS standards such as WS-Policy.

The penultimate stage in the framework is the **Development & Testing phase**. This phase is largely carried out by companies individually, however occasional joint interactions are appreciated for testing, and system verification to previously established requirements. The input to this stage is the agreed systems design specifications (natural language and standards-based) and the service-level agreements. These documents are intended to be used by each individual company (and their personnel) to steer their internal systems implementation.

In the interest of supporting this internal process, the framework builds on current research and suggests the use of guidelines from more detailed and tested approaches such as [9], [8]. In the former work the goal is towards the development process for secure WS. Whereas, the latter article presents a lifecycle methodology that focuses on critical aspects such as application integration, migration from old to new Web services-based processes, and the 'best-fit' ways of implementation which appreciate company constraints, risks, costs and returns on investment. Another benefit to using these particular approaches is that information gathered and produced earlier in BOF4WSS can be reused to quickly complete their initial stages. Such data includes functional, security and QoS requirements, risk assessment data, and business process models. The last step in this phase is to verify that developed systems have achieved the requisite amounts of application-level security. To aid in this, an evaluation is advocated through the use of penetration testing and WS-specific approaches such as those presented in Yu et al. [10].

With the development of this multilayered security solution complete, its upkeep is the next crucial undertaking. BOF4WSS addresses this and other typical monitoring and preservation tasks in the **Maintenance phase**. This stage will involve functional system enhancements, but additionally will stress the continued updating and enforcement of security measures, both in developed systems and the ISS. Cross-enterprise teams both in terms of functional and security aspects are essential to this process. Regarding security specially however, they would be entrusted with monitoring the internal and external environments, and considering new threats, laws, and business requirements, and how these will be included in solution updates.

Having recapped the framework, the next section moves on to consider supporting the transition between two of BOF4WSS' stages, namely Requirements Elicitation and Negotiations phases. Specifically, the section assesses the difficulties incurred in cross-enterprise security negotiations during these stages, and discusses the Model and tool proposed to tackle them, and thereby support phase transition.

III. Supporting BOF4WSS and the Transition Between its Phases

A. The Stage Transition Problem

Sharing, comparing and negotiating on security actions and requirements across companies, even at a high-level, has always been a complex matter. Tiller's work ([7]) gives insight into this issue as he labels the related process, "security mayhem", because of the variety of security aspects (e.g., specific polices, service-level agreements, legal obligations, unique access requirements) to be considered in forming business collaborations. The reality of this problem is underlined by Dynes et al. [11] who set out a research agenda with a core question being: how can a shared vision on risks and security for interacting companies be achieved which appreciates their range of differences?

To investigate the specific issues surrounding stage transition and the negotiation of security actions as they pertain to BOF4WSS, a case scenario was used. This scenario featured companies using the framework during the Requirements Elicitation and Negotiations phases, and especially focused

on how security actions were determined, how these actions were documented/expressed, and how parties compared and negotiated on them. To strengthen the practicality of the scenario, security professionals knowledgeable in external company interactions were interviewed and their input used to guide case development. After defining the case scenario, it was analyzed to identify areas which proved difficult, problematic, or overly tedious for companies. Some of the most prominent areas are discussed below.

- **Understanding the security actions documents of the other companies "as is"**: In the Negotiations phase, companies supply their security actions to their business partners for perusal and discussion. A major difficulty even at this early stage was gaining an appreciation of what exactly companies meant (i.e., a semantic issue) when they outlined a security action or requirement in a few brief, informal statements, often with little justification. Included in this, is the reality that companies may use different terminologies for security actions, associated risks, threats, and vulnerabilities. These problems were further compounded by the variety of techniques (e.g., requirement listings, generic checklists, graphical representations) used by businesses to document their security actions. The core issues at this point therefore link to the *semantic gap* likely to be prevalent across companies, and the *disparity in formats* used to document actions. Both of these aspects resulted in the need for companies to spend considerable time and effort understanding actions and requirements before any negotiations could take place.
- **Understanding the motivation behind other companies' security actions and requirements**: From the summary documentation which constituted companies' security actions and requirements, it was often somewhat challenging for other businesses to determine exactly why that security desire existed. Even if the security situation/risk which the security action intended to address was included in the description, there might have been a plethora of other aspects (e.g., laws and regulations, security policies) considered in the preceding risk assessment that were not specified in the action description. These aspects are important because they provide insight into security actions that form the basis for companies negotiations. As a result of this *incomplete information*, companies usually had to enter further discussions to determine these aspects before making decisions on individual security actions.
- **Comparison of companies' security actions and requirements**: This task entailed parsing through other companies' actions and requirements documents to note and question any existing conflicts across businesses. Included in this task was the implicit or explicit matching of security actions from companies which targeted the same situation or risk. Even in the cases where security actions were classified into groups beforehand, the task of *parsing through documents*, and the various *back-and-forth communications* necessary to match and compare actions even at a basic level, resulted in the consumption of a vast amount of man-hours. An additional issue at this point was ensuring that all aspects motivating security actions (e.g., laws, security policies, contractual obligations) were gathered, documented and readily available for consideration, to support actual comparison and negotiations. Any streamlining of the aforementioned processes would save time, money, and effort for parties.

Having presented some of the core problems discovered from the case analysis, Section III-B outlines the conceptual Solution Model for the system to support stage transition.

B. Solution Model

The Solution Model, shown in Figure 2, contains four components: Security Actions Analysis, Ontology Design, Language Definition and Risk Catalogue Creation. The prime aim of this model is to outline a notional base on which a tool that would actually support the negotiation of security actions across companies, could be implemented. A description of the components is given below.

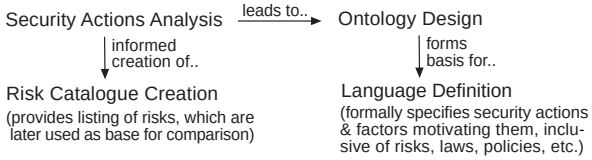


Figure 2. Solution Model

Security Actions Analysis: As a first step to addressing the problems related to the semantic gap and the disparity in formats used to document actions (identified in Section III-A), an in-depth analysis of the security actions and requirements domain was required. This assessment focused on security literature particularly in the security risk management field (as this area was viewed as key to determining security actions), and critically examined how security actions and requirements were derived. From that analysis, common critical factors, especially those that constituted and motivated their derivation were then identified. This component stage's findings allowed for a thorough understanding of that domain, and furnished the foundation for following stages.

Ontology Design: Ontologies are widely known for their ability to specify a shared understanding about a particular domain. In this case, an ontology was used to provide a common understanding of the security actions (and generally, security risk management) domain, based on findings from the Security Actions Analysis stage. Establishing this common semantic bridge was a critical prerequisite in creating the overall solution, when considering how different the terminologies, methods, and influential factors internal to each business were likely to be. It was also important

that the ontology was encompassing, and therefore allowed for an easy semantic mapping of concepts onto it from typical security action determination (or simply, security risk management) methods used by companies. Readers should note that the ontology designed here is high-level and mainly diagrammatic (i.e., there is no formal ontology language). As such, it is more of a communications tool, which can also be built on in future components. An ontology draft, and the Analysis component were previously presented in [12].

Language Definition: Two of the core issues identified in Section III-A center around the numerous formats used for security actions, and the incomplete information initially presented regarding the motivation for those actions. The Language Definition stage addressed these issues by defining a formal language to be used by companies at the end of Requirements Elicitation. The benefit of a formal language as opposed to a shared text-based template, or graphical representation is the automation it would allow; encoded data could now be processed by a machine. This language would enable the formal expression of parties' security actions, and the factors that motivated them (e.g., risks, laws, security policies and so on) in a common format. By having these motivational factors initially included and specified, this negates the need to enter lengthy discussions to determine these aspects later. An XML-based language was preferred to facilitate encoding due to its wide acceptance, XML's platform independence, and the variety of systems support options (numerous APIs for parsing and validation) available. To define the language's syntax, the ontology was an invaluable asset. Aiding in language definition was one of the original purposes of the ontology, as its use ensured that the language was grounded in accepted literature and supported by some common semantics across companies.

Risk Catalogue Creation: To address the problem of matching and comparing security actions across enterprises, emphasis was placed on identifying an aspect which was common to the actions and could be held constant. Therefore, regardless of the divergent security actions for a situation defined by businesses, a common underlying aspect could be used to quickly (or automatically) match these actions. After reviewing the Security Actions Analysis, it was apparent that in a majority of cases, security actions were established to handle or treat some inherent *risk*. The range of security action determination methods used by companies enforced this reality (see work in [12]). To provide the constant base therefore, a shared risks listing/catalogue was instituted and developed. This catalogue contained an updatable, extensive listing of security risks, and was used by companies as a common input to their risk management processes (i.e., the process that identifies, analyzes, evaluates, and decides treatment for the risks). Although businesses used different processes and derived possibly disparate security actions, they maintained a common base in terms of what risks were addressed by a particular action. Once implemented in a system, this common base would allow for the automated matching of security actions from companies, and thus ease the task of matching and comparing actions.

A general idea of how the implemented Solution Model worked towards significantly easing stage transition, is illustrated in Figure 3. In this diagram `Supplier` and `Buyer` are using BOF4WSS for an online business scenario.

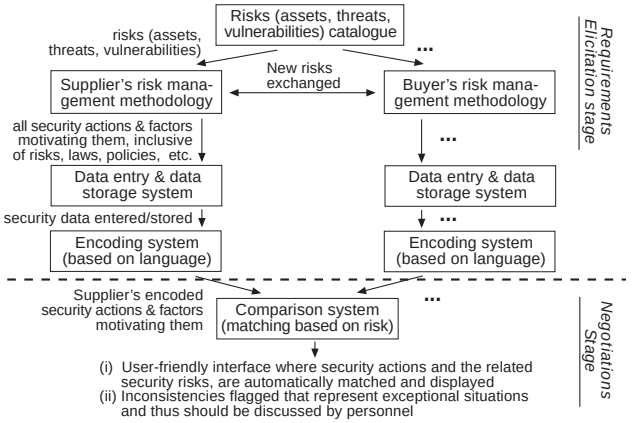


Figure 3. Solution Model in action

A briefly outline is now given on the conceptually implemented model in Figure 3. To begin, risks from the risk catalogue are selected by companies to form input to each entity's risk management methodology (i.e., process to determine security actions and requirements). Once companies determine their individual security actions, these actions and the factors motivating them are transferred into an Encoding system and marked up into the XML-based language defined. When businesses meet in BOF4WSS' Negotiations stage, the encoded documents are then passed to a Comparison system that matches companies' security actions based on the underlying risks they address. Currently, the output of the Comparison system focuses on (i) a user-friendly interface where security actions (supported by related risks, and motivational factors) are automatically matched and displayed, and (ii) flagging of any inconsistencies identified for follow-up by personnel. A noteworthy point is that the Solution Model and resulting tool are especially geared towards *shared* risks faced by entities. Therefore in some regards, emphasis is placed on the shared risks where companies have to agree on how they will be treated i.e., the type of security action (e.g., mitigation, transference, acceptance, avoidance), and actual action to apply. Section III-C formally introduces the tool which embodies the Encoding and Comparison systems above. This is the Security Actions Specification and Comparison System, hereafter SASaCS.

C. SASaCS Tool

1) Overview: The SASaCS tool represents the culmination of this work, in that, it is the software implementation of the Solution Model. SASaCS consists of all the practical components necessary to support the presentation, sharing, comparison and negotiation of security actions across companies. As a result of its tight coupling with the Solution Model, the general process outlined at the end of Section III-B applies to the tool as well. In Section III-C therefore, we provide more detail on the tool by discussing three of its features, the Data Entry interface, Comparison System report output, and the Encoding system (XML language). These aspects were chosen because they allow novel parts of SASaCS to be highlighted, and set the platform for evaluation in Section IV.

Once companies have conducted their risk management activities (which are informed initially to some degree, by the shared risk catalogue) and produced their individual security actions, the next task is transferring them into (their locally installed copy of) the SASaCS tool. This is handled by the Data entry and storage system. This system, shown in Figure 3, provides a set of simple, intuitive screens for users to input their security related data (e.g., risks, security actions and factors motivating them) and have it stored to a back-end tool database. To ease usability, the tool also allows the direct referencing and selection of risks from the risk catalogue, that initially factored into the company's risk management activities. Therefore, users can look-up risks from the catalogue, apply them to the current project/collaboration, and then annotate them, or otherwise use them as they see fit (e.g., input' risk priority levels, associate them with a security action, and so on).

As SASaCS is based on the ontology designed, its data entry screens benefit from the unambiguous definition of concepts (such as risk, risk level, and so on) prevalent with the ontology. The ontology diagram itself and its documentation also are useful in assisting users understanding of concepts, and linking data entry fields to output from their risk management methodologies. In addition to having data fields mirroring the basic concepts from the ontology, the Data entry interface defines a number of other fields to allow companies to add more detail on relevant aspects such as company-specific risk descriptions, justifications of risk levels, annotations regarding treatments of risks, treatment coverage levels, and security requirements. Figure 4 shows a screenshot of the security action (or in other terms, risk treatment action) data entry screen in SASaCS.

After each enterprise has saved their security- and risk-related data to the tool, the following step is encoding that data in preparation for inter-company negotiations. The Encoding system (also installed locally) facilitates this by pulling data from the tool database, marking it up in the XML-based language discussed previously, and outputting a document with the encoded data. When companies meet for negotiations therefore, (i) they use the same format to express security actions/requirements, which is also machine-processable; (ii) there is a shared understanding of the security- and risk-related concepts, promoted by the common ontology and highly supportive tool data entry screens; (iii) information is more complete as factors motivating security actions should initially have been supplied; and (iv) because encoded data (particularly security actions) includes references to risks in the risk catalogue, there are commonalities across companies' documents. The Comparison system uses these commonalities to automatically match security actions/requirements that treat the same shared risks.

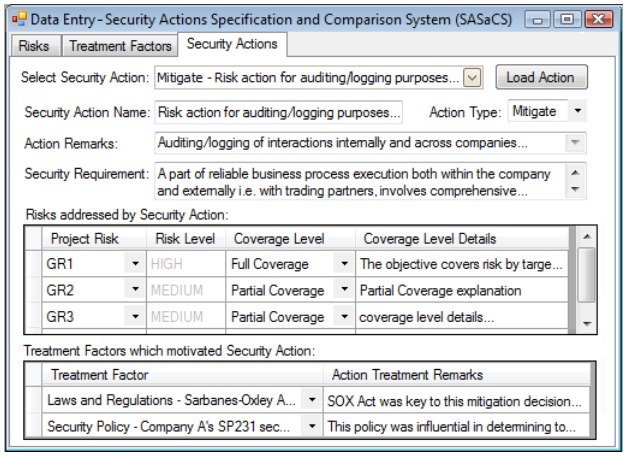


Figure 4. Security action data entry screenshot

As an example of the process mentioned above, let us assume two companies, `Supplier` and `Buyer`. Furthermore assume Figure 4 is a screenshot taken of SASaCS running at `Supplier`. There therefore exists a mitigation action formulated by `Supplier` to handle three risks, *GR1*, *GR2* and *GR3*. Reasons for their decision are listed in the treatment factors subscreen. At `Buyer`, assume that personnel only consider risk *GR1* and *GR3*; *GR1* they opt to mitigate, and *GR3* they choose to accept due to limited a security budget. By having all this information supplied in the system initially, when parties meet for negotiations, SASaCS can be used to quickly assist in various important tasks. One such task is automatically matching the disparate security actions of `Supplier` and `Buyer` based on underlying risks. Figure 5, which displays output from the Comparison system based on data above, exemplifies this. Here, companies are immediately notified of conflicting security actions (for example, in the treatment *GR3*), and situations where some entities do not address risks at all (in the case of *GR2*, by company `Buyer`). Additionally, businesses are instantly shown key reasons which motivated each company's particular security action decision (by way of treatment factors).

Streamlining these, at times simple tasks, can significantly reduce the time and effort needed by companies during the

Figure 5. Security action report output screenshot

initial stages of BOF4WSS negotiations. In the next section, we examine the encoding aspect more by presenting the XML-based language defined. For ease of reference, this language is called SADML, or Security Actions Definition Markup Language.

2) The Language: The structure of SADML was conceived to mirror the knowledge captured in the ontology (largely defined in [12]). As such, various ontology's concepts are represented as XML elements/tags. To comply with XML's hierarchical nature, it was necessary to define a sensible hierarchy of elements. Furthermore, this structure would need to accommodate one-to-many relationships across elements (for example, if a single law motivates/supports multiple security actions, this should be appreciated). Considering these and a few other salient aspects, SADML's syntax was defined. A snippet of the SADML format representing the information in Figure 4 is presented below; the + sign indicates additional data which is not displayed here for space reasons. The core language is described in the schema, indicated by `urn:risksx-schema` in the snippet.

```
<needsBase xmlns="urn:risksx-schema" ... >
  <mitigationActions>
   <mitigationAction>
    <name>Risk action for auditing/logging...</name>
    <details>Auditing/logging of interactions...</details>
    <risks>
+     <risk id="GR1">
+     <risk id="GR2">
+     <risk id="GR3">
    </risks>
    <lawAndRegRefs><lawAndRegRef idref="LR22">
      <relationToRiskAction>SOX Act was key to this miti-
       gation decision based on...</relationToRiskAction>
    </lawAndRegRef></lawAndRegRefs>
+   <securityPolicyRefs>
+   <securityRequirementRefs>
   </mitigationAction>
  </mitigationActions>
+ <acceptanceActions>
  <transferenceActions /> <!-- No actions defined -->
  <avoidanceActions /> <!-- No actions defined -->
+ <lawsAndRegs>
+ <securityPolicies>
+ <securityRequirements>
 </needsBase>
```

As can be seen above, `needsBase` is the root element and its sub-elements encompass the four general types of security action, and the main factors identified which motivate them. In practice, SADML groups risks by the *type* of security action (e.g., mitigation, or `<mitigationActions>`) which addresses them, and then the exact written action (e.g., `<mitigationAction>`) defined by a company. Because one security action can address many risks, each action has a `<risks>` element that lists the risks addressed. The elements suffixed with 'Refs' are used to indicate that existing motivational factors, for example laws and regulations (`<lawsAndRegs>`), influenced the treatment of a risk. `<securityRequirementRefs>` is the exception, in that it references security requirements (`<securityRequirements>`) that detail security actions. SADML's structure proposes one way to define security actions, risks and motivational factors, and does not intend to be a panacea in itself.

The novelty of SADML is rooted in the unique business perspective it takes on risks and security actions, which aims to (i) maintain a strong practical foundation (by mirroring the ontology designed) and (ii) place security, at least initially, at a level that understandable to security professionals and business-based decision makers (often the budget holders) alike. Next we cover existing work related to the Model and tool.

D. Related Work

In [13], authors assessed similar disparity problems to the Solution Model, particularly in communicating security requirements. They proposed a framework for formally specifying requirements and detecting conflicts amongst collaborating parties. The difference between that research and our work is in the layers which are targeted; the Solution Model supports high-level security negotiations for businesses, whereas Yau and Chen [13] consider low-level security requirements (and by extension, only risk mitigation), and formal rules and algorithms for requirements refinement. Their approach therefore is not actually concentrated on the problem which our work emphasises.

Apart from the related literature on the ontology previously presented in [12], the only other area with similar work is the XML-based language defined. In research and industry there have been a plethora of security languages covering from access control (e.g., XACML), to identity management (e.g., SAML). The most relevant to our work is the Enterprise Security Requirement Markup Language (ESRML) [14]. This language is comparable to SADML because it emphasizes the higher layers of security, and the sharing and exchanging the enterprise security information across companies for business purposes. The shortcomings of ESRML in terms of this work however are its lack of emphasis on factors which significantly influence or drive security actions (e.g., regulations, constraints), and its concentration on risk mitigation as opposed to explicitly appreciating other ways to treat risks.

Having now covered the framework and the Solution Model and tool, Section IV reports on the interview-based evaluation conducted on these proposals. This evaluation and its findings form the key novel contributions of this paper.

IV. Evaluation and Findings

A. Evaluation Method

To evaluate BOF4WSS and the Solution Model and tool, a standard structure of research was followed. This included the definition of areas of interest and then the collection and analysis of relevant data to assess these areas. Rigid hypotheses were not preferred because this evaluation does not seek to thoroughly prove or disprove formal theory. Instead, the aim is to establish whether the information gathered supports the areas and proposals assessed, and if so, the degrees of support arising from the data gathered.

There were two core *areas* to be investigated for support in this evaluation. First was to investigate whether the framework proposed is an applicable, practical proposal which would aid businesses in reaching requisite levels of enhanced inter-organizational security and trust. And secondly, to examine if the Solution Model and tool provide a viable process to greatly support transition between the Requirements Elicitation and Negotiation phases of the proposed framework.

To study these areas, a qualitative research strategy was chosen in which digitally-recorded, semi-structured interviews were employed. The interview data gathering technique was preferred as it allowed for a detailed study into the field and the gathering of descriptive, insightful data for analysis [15]. Semi-structured interviews enhanced this process because they allowed for a mixture of structure and flexibility in questions asked. Therefore, in addition to asking planned questions which directly related to the areas above, other interesting and associated observations could be explored.

To ensure the interview questions were clear and appropriate, pilots were used to refine them initially. Also, in the interest of gaining the highest quality feedback, interviewees were sent general documentation on the models at least a week before the interview. This allowed them time to review the proposals and gather their thoughts before the meeting.

As was mentioned, the target group for interviewees consisted of industry-based security professionals. To narrow this further, purposive sampling [16] (which is the use of special knowledge to select appropriate subjects) was applied. Within this general group therefore, individuals were selected that showed a good experience (demonstrated by job roles, certifications, qualifications, and past project involvements) in the following pertinent fields: Web services technology, e-business and online business paradigms, security risk management, information assurance, security architectures, and cross-enterprise interactions.

Specifically, the interviewee selection process consisted of directly contacting persons with demonstrated experience (identified from company Web sites and/or articles published), and also using the author's contacts within companies to help identify other relevant professionals. It should be noted that no special incentives for participation were offered and interviewees participated based on their own free will. This targeted selection technique was adopted as opposed to more statistically random or quasi-random techniques, to ensure that persons selected had a good degree of requisite experience and specialized knowledge.

Additionally, because the emphasis was on gathering in-depth information rather than surface-level data from as many persons possible, only five professionals were interviewed. These professionals however had a total of 48 years experience in the security field. This small sample size allowed for a manageable, yet very detailed amount of expert feedback to be gathered in the, on average, two-hour long interviews. Small sample sizes, greater depth of information, and a focus on narrative data all are key characteristics of purposive sampling [17]. Known limitations of this sampling technique however include possible bias in interviewee selection, and lack of wide generalizability of findings [16]. As there was no relation between subjects and the interviewer and as subjects were selected based only on demonstrated experience and no knowledge of their personal opinions, bias was not viewed as a serious limitation here. Furthermore, wide and conclusive generalizations are not the goals of this evaluation but rather to gain some insight into the use of research proposals. This wisdom might also then be applied in the next stage of evaluation, i.e., the case study.

Therefore, although there are noteworthy limitations of purposive sampling, the benefits possible with the technique were seen to outweigh the drawbacks in this case. This is especially considering the resource and time constraints on this project, and great amount of time taken even to set up interviews with the five subjects chosen. (Common

issues faced were the busyness and hectic schedules of professionals, coupled with the need for companies' legal departments to be involved to consider and approve the interviewee's participation.) Finally, to encourage honest and detailed feedback, the interviewees were told that their identities would be kept anonymous. This also avoided any more possible legal complications with their companies.

The overall goal of the interview process therefore, was to present BOF4WSS and the Solution Model (particularly, core characteristics, possible areas/scenarios of contention, novel aspects of them), and attain a real-life, expert opinion and in-depth insights. This feedback would delve into the applicability (how suitable are the models for the situations and problems they target, what might the response from companies be) and strength (how well, if at all, are the problems addressed by models, what are their benefits and shortcomings) of the proposals based on security professionals' real-world experiences.

Having conducted the interviews, recordings were then transcribed. To analyze the data collected, the content analysis [16] data analysis technique was then applied. This provided a standard method to code, organize, and index the transcribed interviews. Furthermore, it allowed for easy data retrieval, pattern identification and review, and basic counting to note any relevant quantitative observations [16]. A blend of deductive and inductive approaches to identifying themes in the data was favoured. This enabled themes to be identified which focused on the investigating of the areas for support (deductive) but also common themes that arose from data that were not conceived prior (inductive).

With the research process outlined, the next section concentrates on the presentation and analysis of the research findings. This research interweaves the findings and analysis stages because it was felt that this would allow for a rich but also concise discussion. Berg [16] supports the viability of this combined option especially when compiling reports based on qualitative data.

B. BOF4WSS

The first *area* to be investigated centres around whether the framework proposed is an applicable, practical proposal which would aid collaborating businesses in achieving desired levels of enhanced inter-organizational security and trust. To examine this, questions to interviewees concentrated on core principles and novel aspects of the framework which specifically aimed at addressing the outstanding research problems. Four *themes* have been identified in which to present and analyze the data gathered.

The themes consider: (i) the framework's emphasis on a highly collaborative approach to inter-organizational security, particularly where WS is concerned; (ii) the reality that BOF4WSS is detailed and at times prescriptive; (iii) the merit of the framework's focus on higher layers (business-level for example) of security in WS-based cross-enterprise interactions; and (iv) the use of the Interaction Security Strategy (ISS) as a comprehensive security management structure, that could also foster trust across partners.

Using interviewees' feedback, the themes are assessed in terms of their use and/or strength, and application. After theme analysis, an additional section is presented including interviewees general comments on the framework, before briefly summarizing the assessment thus far. In the presentation below, fictitious names are used for interviewees. This respects their anonymity while also allowing for a more vivid presentation of findings.

1) BOF4WSS and its highly collaborative approach: BOF4WSS emphasizes a highly collaborative approach to cross-enterprise security. This high degree of collaboration (manifested in dedication to working together, a good degree of information sharing, various meetings, and other time and investment commitments) was conceived specifically to address the shortcomings stemming from the isolated and individualistic approaches to securing e-businesses which use WS. Noting the amount of stress the framework places on this topic, it was chosen as one of the areas to evaluate within the interviews. The aim being to determine whether highly collaborative approaches such as the framework, might provide more adequate solutions for WS-based e-business interactions, as opposed to more individualistic approaches. The subsequent aim would be to then identify how applicable and practical such approaches are.

In response to questions posed regarding high degrees of collaboration as opposed to individual approaches to security, all professionals expressed that these types of approaches were preferred and yielded better security solutions. Interviewees indicated that solutions were likely to be more appropriate, skills and knowledge could be pooled, and finally systems could be designed and integrated more securely. This favourable opinion was upheld by professionals when questioned about BOF4WSS and its collaborative efforts towards security as well. An interesting point put forward by one professional was that collaboration (especially initial meetings and willingness to work together) enabled him to be able to determine whether or not other companies were really committed to interactions and security or not. Collaboration was therefore being used as a tool to learn about potential partners and even their security postures *before* entering fully into business interactions with them.

Considering collaboration in the context of WS and BOF4WSS, John, a security professional of 10 years working for a leading international IT and consultancy services company, noted that collaboration is essential and needed at all levels (business, legal, and technical agreements). Continuing the Security Architect said, "... particularly with Web services, it has great promise but it's only going to work with that sort of collaboration". This view hints to an importance of an increased amount of collaboration, even within the technology-driven WS world. Detailed feedback

from other interviewees supported the importance of collaboration between companies in achieving inter-organizational security. Existing case study data (see Todd et al. [18]) can also be referenced to see a glimpse of benefits of collaboration.

Even though supporters of collaboration, two professionals warned that it was important for businesses to maintain some degree of individuality (in terms of self-defense capabilities), or at least some safety net features (contract- or technical-based) within collaborations. These would protect individual companies if their partners inadvertently or intentionally became rogue. This point acts as a reminder that collaborative security approaches should not only focus on protecting the group of entities, but also protecting individual enterprises from risks of being in the collaboration (for example, see those raised by Baker et al. [19]).

Having looked at the use of highly collaborative approaches in building cross-enterprise security solutions, the next step was to assess the application and practicality of such approaches, and the framework in particular. From the feedback received, two opposing views were apparent. Three professionals regarded high degrees of collaboration across companies as difficult to attain, whereas the others saw it as "quite practical", and not "too big a barrier". The main proponent for the former perspective was Mark, an Information Assurance manager in a global telecommunications and consultancy firm.

Drawing upon his 20 years in the security field, Mark stressed that collaboration was beneficial to have, but very difficult to attain. Additionally, making persons communicate, work together, and readily share information (which are key activities in a collaborative process such as BOF4WSS) were not easy tasks. Prime reasons cited centred around stakeholder-related issues, particularly the likely problems incurred when meshing teams from different companies with possibly different perspectives, processes, systems, and organizational cultures. These issues are supported by literature in [20], [21].

Interestingly, John also showed an appreciation for the collaboration difficulties mentioned above but did not view them as too much of a barrier. Instead he noted, "yes it is intensive and costly to some extent and I think that's the only way to be really successful". In spite of these difficulties therefore, in his opinion, these approaches were not only practical but a necessity for success with security. Literature could be seen to support this 'security success via collaboration' perspective but primarily in closely knit business partnerships such as the extended enterprise (see Dynes et al. [11]).

Considering BOF4WSS in more detail, additional notable difficulties were identified by subjects relating to complexities in stakeholder arrangement and management (getting the right people together at the right time from across companies) and cross-border collaboration issues (in essence, normal collaboration issues exacerbated by ranges of cultures and perspectives) if/when the framework was applied internationally.

Speaking objectively, the aspects mentioned were somewhat overlooked in our creation of BOF4WSS due to the assumption that shared business aims, and goals for security would drive and support collaboration. When this assumption was put to subjects, some respondents agreed that shared aims would help. However, they also expressed that there would need to be strong, mutually understood benefits for all companies, degrees of fairness ("Nobody wants to be the weak partner", Mark stated), and executive sponsorship from businesses. High-value projects and situations where there was positive history (and existing trust) between companies were also cited as scenarios in which high degrees of collaboration would be more practical. All of these driving factors would have implications for BOF4WSS and indicate situations in which it might be best used.

As a brief summary to the information above, there was some consensus that the high degree of collaboration advocated by BOF4WSS would lead to a more adequate security solution for cross-enterprise interactions. According to the data however, its applicability may be limited (or at least, best suited) to business scenarios where either there is a strong commitment to businesses goals (and security is seen as an enabler to those), a substantial degree of executive sponsorship, they are high-value projects (amount stood to be gained or loss, motivated need to do whatever necessary to get job done), or there is existing trust between companies. The first two of these were previously mentioned in [2], [3] as criteria for businesses adopting BOF4WSS. Conversely, the need for positive history and some degree of existing trust between companies was not envisaged before as a prerequisite to adoption. This was a significant finding as it suggested that even though the framework was aimed at building trust across partners, some history or trust should already exist.

2) Detailed and at times prescriptive framework: In seeking to create a comprehensive security-focused methodology (which supported companies from the planning to maintenance of cross-enterprise interactions using WS), a central objective of BOF4WSS was to provide detailed, and occasionally prescriptive guidance. This guidance included the activities that might and should be conducted, possible ways in which they could be conducted, and their pertinence to attaining desired levels of layered security within the foreseen cross-enterprise interactions. With appreciation of the detailed level of guidance and the possibility that it might not be well received by companies, it was chosen for assessment in the interviews. The objective was to ascertain its usefulness and applicability in aiding the creation of a security solution.

From an analysis of the data, it was seen that a majority of professionals found the detail in BOF4WSS (exemplified

through the presentation of the framework's phases) of benefit to companies, and felt that enterprises would and should be open to it. Some of the benefits they quoted included the fact that detail would force people to consider all the factors, and give structured ways—especially for inexperienced persons—to solve security problems. Another benefit seen in the framework was the visibility and ability to audit, it would bring to all aspects of the cross-enterprise scenario. According to Matthew, head of Information Security and Risk Management at a higher educational institution, "An audit department would absolutely love this". This was stated because the framework would define a structure that audit departments, even though not security specialist, could follow and use to track and compare projects and other company interactions. It should be noted that Matthew has worked in other businesses in IT security roles previously. He also expressed that issues in core business and education (at his institution's level) were very similar.

The main warning placed on the framework by professionals was that it should be wary of being detailed and prescriptive to the extent that companies were not allowed to adapt parts to the nature/culture of their enterprise. This could relate to tools, specific techniques, or constituent methodologies. As is seen to some extent in Section II and largely in [3] however, the framework appreciates these issues and either provides a set of options (such as a listing of risk management methods to determine security needs), or relies on industry standards and best practices (including use of ISO/IEC 27000 for security or UML for modelling).

From the findings above therefore, it can be concluded that the detail provided by BOF4WSS should be useful to businesses and more of an advantage than a hindrance. This would not only apply to persons and businesses that lack experience in dealing with security issues in WS interactions within an e-business context, but also to entities seeking to have a framework to maintain structure, consistency and visibility in the overall process.

3) Appreciation of higher layers of security in cross-enterprise interactions: Another main aim of BOF4WSS is to emphasize holistic security solutions. Holism is used to refer to an all-encompassing approach that considers technologies, policies, processes, methodologies and best practices for security. This aim specially attempts to combat the overly reliant focus on technical mechanisms for security discussed in Section II. The purpose of this section therefore is to evaluate that aim and its merit in the context of cross-enterprise WS interactions.

Commenting on the data gathered, all interviewees displayed an appreciation of high levels of security and echoed the sentiment that technical approaches alone were insufficient. This finding therefore supported the framework's charter and literature in Singhal et al. [22] which highlighted the need for the higher layer of security with WS.

Speaking on this topic, John remarked that the challenge found in business today was achieving this higher level of engagement in projects, specially business ownership, and business and ICT alignment. Technology-level integration was not a problem but rather getting the engagement, involvement, and buy-in for projects at the higher business levels, security-related and otherwise. Lack of these higher level aspects, he noted, were the reasons many projects failed or stalled. Considering this challenge in terms of BOF4WSS, there is a focus on the higher layer, however no special mechanisms of encouragement to achieve it are provided. In the framework design it was envisaged that there would be a top-down drive for projects and therefore efforts were concentrated on supplying guidance for the necessary processes.

An additional concern lodged by two professionals was that even though the higher layer of security was important, the translation and implementation of these higher aspects to lower levels were equally important and not to be neglected. Paul, a Senior Security Researcher at another well-known global IT company, warned that various things get lost in translation and imperfect implementations. This can be to some extent supported by difficulties highlighted in [23], [24]. Furthermore, Paul stated that, "you cannot solve problems at the highest level, that's the thing, you do have to come down to the lowest level". As a result of these factors, he highlighted that it was key that security go through the entire process and the framework should maintain a balance between higher and lower layers to security, and not overly emphasis either. This was an accepted perspective in BOF4WSS as it aims for holistic security.

Continuing the assessment on the merit of higher layers of security, the next question to interviewees centred on trust, and whether this layer (and the activities therein such as jointly defining policies, agreeing on process for security, meetings and so on) in BOF4WSS might lead to increased trust across entities and their personnel. In response to this, a majority of professionals agreed on the likelihood of increased trust resulting. Common rationales presented linked to time spent together and commitment towards security that, once present, would be demonstrated to partners. Both of these could lead to relationship building, which then may lead to trust. Todd et al. [18] is one documented real-world scenario where high-level activities such as joint risk assessments, "proved to be the foundation upon which mutual trust between the security communities ... has been built" [18].

Mark was the least enthusiastic about the higher layer naturally achieving trust as he felt that trust was a very complex and difficult thing to attain—a view supported by Van Slyke and Bélanger [25]. This he attributed to human factors and the difficulty in predicting human behaviour. Aside from this however, respondents' feedback supported the possibility of increased trust across business partners.

4) Use of the Interaction Security Strategy (ISS): The Interaction Security Strategy (ISS) is one of the more novel parts of BOF4WSS, in that it seeks to create and apply a cross-enterprise management structure not found to be used in practice. The first question to interviewees therefore was to gather their opinion on this strategy in terms of security and trust. Another point of interest was how the strategy compared to existing approaches, particularly contracts, as these seemed to be the main agreements structure used today by companies.

The feedback gathered indicated that a majority of security professionals felt that the ISS was a valid and useful approach for cross-enterprise security and trust. Only Luke, a Senior Security Researcher with 4 years experience, disagreed as he was not sure about ISS positioning in the framework's process flow, or the level of security present in the ISS; he regarded it as too detailed.

One intriguing finding was that even though legal contracts formed the main agreements mechanism across companies, they were reported to cover security only very generally. For example, if in the UK or EU, they might only very briefly reference the Data Protection Act. Drawing on his 10 years experience, Matthew highlighted that contracts are not likely to cover security policies, continuity planning, or even ISO/IEC 27000 best practices. He emphasized that it was therefore important to have an extra layer of security (similar to the ISS) in place. Generally supporting this point, a 2010 survey [26] has highlighted that roughly 40% of large business respondents do not ensure that their contracts with third party providers include security provisions. This is a telling aspect in terms of contracts and their lack of focus on security.

Additional advantages of the ISS identified by some interviewees linked to the flexibility it would allow, and the pragmatic, actionable structure it provided over contracts. Contracts were seen to be very specific, hard to follow, and often expressed in legal jargon. The key stipulation made by subjects however, was that the ISS was always in line with the contracts. This, they stated, would ensure synergy in agreements. In general therefore, professionals' feedback above is seen to support the ISS as a key tool in creating and instilling a cross-enterprise security solution. This would enhance the practical security provided today and support agreements in contracts.

The second question related to the ISS concentrated on its use as a mechanism to foster trust across businesses. Trust was hoped to be achieved by making security approaches (pertaining to the scenario) more predictable and transparent (these being two key attributes of trust [27], [25], [28]). From the resulting interview data, a consensus was apparent as professionals all regarded the ISS as likely to foster trust. Reasons supplied included the clear guidance to companies, and the ownership and understanding it supplied personnel with, considering that they aided in its creation. Both of these aspects link with intended goals of ISS. John's support for the ISS in this regard was motivated by its charter towards a joint security posture, something that he felt was more conducive to trust, rather than the "us and them" mentality he saw in some businesses today. This opinion can be related to collaboration in general and the reality that some entities might not be willing to collaborate to this extent.

The other salient view on the ISS and trust was held by Mark. He expressed the view that,"[the ISS] probably fosters trust in that it takes away distrust ... What you'd certainly find is that one of the major hurdles is getting over the distrust, doesn't mean that you've actually got trust once you've got over that". This view, albeit a solitary one in the context of respondents, highlights the precarious nature of trust and possible difficulty in gaining it across persons and enterprises. In general however, the ISS is seen to positively aid in this venture and provide a structure that could enhance currently used mechanisms.

5) General thoughts on the framework: With the framework's core principles and novel aspects assessed, the next three paragraphs highlight other noteworthy feedback (based on consensus, ideas related to research literature, or simply practicality) given by interviewees.

One view that arose with respect to security frameworks and methodologies generally, was the inherent difficulty they faced in balancing complexity and being comprehensive, with making them useful and consumable by businesses. John aptly summarizes this opinion in his remark, "getting the balance right is so important where it's rigorous enough to add value and to make sense, make the process more structured, and at the right level but not so verbose that it's not useful". He further stated that even though the real proof would be in the adoption of the BOF4WSS, to him, it looked okay and seemed "light enough ... to be useful".

Another intriguing point which surfaced was that BOF4WSS did not appear to be specially suited to medium-to-high security or trust industries or business scenarios. Instead interviewees felt that it was generic and according to Matthew, "would be good across the board". This perspective was of interest because the framework was originally targeted at businesses and scenarios that emphasize trust and medium-to-high levels of security (see [2], [3]). These cases were chosen as they were seen to justify the significant effort and resources needed to adopt and use BOF4WSS. Based on the data collected however, the framework might have wider scenario applications, subject to limitations from other findings.

The final significant point relates to framework applicability again, but more from a higher perspective. In considering the application of BOF4WSS to scenarios, Paul expressed that asymmetries (whether due to size or bargaining power) in the market might limit the framework's use. This was because asymmetries lead to some enterprises looking to

develop solutions (usually individually) to service as many generic customers as possible. This was as opposed to focusing on one-to-one collaborations and individual partner requirements (such as purported by the framework). Albeit a notion only mentioned by one professional, the collaborative nature of BOF4WSS might suggest that it is better suited for symmetric-type interactions. These are interactions where each party has an influence, and party-to-party negotiations, design, and development is expected.

6) Summarizing framework analysis: Having presented and analyzed the main findings related to the framework, below these are briefly summarized and used to investigate the degree of support for the *area* highlighted at the beginning of Section IV-B.

The first area was the most debatable and investigated the high degree of collaboration desired by the framework. Based on the analysis in that section, collaboration was likely to lead to more adequate and thereby enhanced solutions than those possible with individual or isolated approaches to security. Additionally, it was also concluded that BOF4WSS (and to some extent, highly collaborative approaches in general) may be better suited to certain business situations and scenarios because of their nature (see the collaboration theme discussion for details). These findings strongly support the area being investigated, but limit the target scenarios of the framework.

Considering the level of detail provided by BOF4WSS, a majority of interviewers saw this as a benefit to companies which would, and should be welcomed. This was assuming that it allowed some degree of flexibility, which it can be said that BOF4WSS does (through the provision of various tool/technique options). Cited benefits of the framework included forcing companies to consider all the factors, aiding inexperienced persons (in what is arguably still a relatively immature field in terms of WS use for supporting complex business processes), and creating a level of visibility and ability to audit, for cross-enterprise development and subsequent interactions. These aspects can all be seen to enhance current security approaches and therefore provide good support for the area studied.

Reflecting on the appreciation for higher layers of security in the context of WS in e-business, data showed a consensus in their merit and value within the overall security approach and solution. The main concern identified at this stage related to getting the necessary level of engagement, at what is essentially the business layer within companies. This is a problem not covered by the framework as it was assumed the necessary top-down drive for projects already existed. This top-down drive would be present in the applicable scenarios suited for BOF4WSS, highlighted in the sections above.

On the topic of trust, a majority of positive interviewee feedback acted to further support the framework's appreciation of, and concentration on this higher layer. To recap, this layer involved getting companies together to interact, collaborate, and discuss and plan interactions security. Generally, these findings are therefore considered to provide a noteworthy degree of support for the area being investigated, both in terms of security and trust.

The ISS is in many ways a specialization of the higher layer security approach covered above, and interviewees also saw it as a useful approach in terms of cross-enterprise security. Its importance was accentuated particularly because there seemed to be no standard overarching management or guidance structure for businesses which pertained to security. Contracts were referenced, but it is known that these documents do not contain detail on security nor do they place it in an actionable language and context. Furthermore, findings indicated that trust between companies was likely to be fostered by the ISS. Interviewees linked this to the transparency and clear guidance for companies, and ownership and understanding implied as companies would have aided in the creation of the ISS. In terms of the area for support, the novelty in the ISS was seen to add to current approaches both in terms of security, and possibly also regarding trust.

Based on the preceding paragraphs and sections, it can be concluded that in the context of this evaluation, there is significant support for the framework. This support is with respect to providing an applicable and practical approach to enable businesses to reach requisite levels of enhanced cross-enterprise security and trust. Critically speaking, the majority of support for the use and viability of the framework, relates to business scenarios where there is either: a strong commitment to businesses goals; a great degree of executive sponsorship; they are high-value projects (and this value drives the need to do whatever necessary to complete the task properly); there is history and existing trust between companies; and there is symmetry in business interactions. Based on these characteristics and predefined target areas for the framework as defined in [2], [3], specific candidate companies that should benefit most from BOF4WSS adoption are:

- Large companies with smaller units (or subsidiaries) seeking to streamline online interactions using WS between these smaller units — As part of the same company, executive sponsorship and strong commitment from parent units would be a strong driver for smaller units to collaborate and bring interactions to fruition. These units would be focused towards symmetric collaboration therefore there would be the need for both parties to engage in context-specific negotiations, design, customization, and development. Also, assuming history between these units (given that it is the same company) there will already be a foundation of trust that can be exploited and built on.
- Partners in an extended enterprise setting, for example e-supply chains — Research in extended enterprises aided in the construction of this framework and a number of the criteria listed above meshes with needs

in these types of business networks. As trust is already a key prerequisite in extended enterprises [27], if a group of businesses in such a network desired to switch from proprietary integration formats to WS for cross-enterprise interactions, BOF4WSS would be very useful. The long-term nature of these networks and strong commitment towards a shared goal and mutual benefits also support the framework's use. Furthermore, because these businesses tend to already be collaborators at the strategic and business level, collaborations in security using BOF4WSS would be a natural next step to protect inter-organizational interactions and individual enterprises. Symmetric interaction would also apply.

– Small and medium-sized enterprises (SMEs) seeking to build long-term partnerships — This relates in particular to small and medium-sized companies with past history, a strong commitment to partnerships, sustained symmetric interactions, and the desire to achieve shared business goals realized using WS. BOF4WSS would be of great applicability to these type of companies for two reasons. First, because there might be a lack of expertise and experience, the framework's detailed guidance would be very useful. Second, as there are less stakeholders, stakeholder arrangement and management should be less of a problem. To justify the time and resources necessary by BOF4WSS, long-term alliances are likely to be the most practical scenarios. In such situations companies can see their investment yielding returns in the long-term.

The next section presents the findings and analysis conducted regarding the Solution Model and tool.

C. The Solution Model and Tool

In this section, the second core *area* is examined to determine whether the findings support it, and if so, to what extent. Specifically, this involves an investigation into whether the Solution Model and resulting tool provide a viable process to support transition between the Requirements Elicitation and Negotiation phases of the framework. Similar to the evaluation of BOF4WSS above, questions to interviewees assessed novel characteristics and core precepts of the Model and tool.

For the presentation and analysis of data, four *themes* have been chosen. These include: (i) opinions on transition problems highlighted; (ii) the premise that risks drive security actions and requirements; (iii) the likelihood of business partners sharing detailed information on common risks and their intended treatments; and (iv) the ultimate use of the Model and tool. Data within these themes is analyzed with respect to its application and scope. As with Section IV-B, there is a final section that summarizes the conclusions from the analysis completed.

1) Opinions on transition problems highlighted: The charter of the Solution Model was to address the transition problems that companies were likely to encounter in moving from the Requirements Elicitation to Negotiation phases in the framework. These problems were identified based on an informed case scenario and relevant research literature. Considering their importance as a driving factor for the Model however, this theme assesses the issues again with the goal of determining exactly how serious they might be from professionals' perspectives.

Commenting on the feedback received, all but one security professional—i.e., Luke—agreed with the transition issues highlighted. In response, Luke said he was unsure whether security would be considered at what he considered, an early stage in negotiations. In cases where there was agreement, professionals concurred with all of the transitional problems (such as semantics issues, difficulties understanding motivation for actions, and the arduous task of comparing and negotiating actions), and substantiated their opinions by drawing on past experiences.

In terms of semantics issues during phase transition, John stressed the importance of spending time initially agreeing on terminology in projects, as words in the security domain are often misused. Paul and Matthew were two of the main proponents supporting the reality of disparity in formats of security actions and requirements. Relating to this, Matthew stated, "there are companies that might have a basic statement, they might have a graphical representation, they might have a few bits and pieces and in my experience actually getting those to marry together initially, is one of the hurdles you do have to get over". These aspects can be compared to the security mayhem discussed by Tiller [7].

One of the most interesting findings in the data related to the motivation behind security actions and requirements. On this topic, John noted that in addition to partners not supplying (or supplying little) motivational information initially, if they were asked to justify actions at a subsequent stage, they did not always have good reasons to support their security actions. He explained that in some situations where standard security actions (such as reused action lists, or generic security checklists) were provided by companies, the original meaning might have be lost, or the security landscape might have changed. Therefore in addition to the problems associated with businesses not communicating the motivation behind security actions, the reality exists that companies themselves might not be clear about reasons for their actions. This adds an extra level of complexity and discussions as companies meet in the Negotiations phase.

Another noteworthy observation from the data was that personnel involved in cross-enterprise negotiations may not always have a security background—they may be business-oriented persons for example. Matthew felt that some personnel have basic knowledge of security aspects but because they lacked core knowledge and experience in security, this tended to prolong the negotiations process. This is important because it highlights that even though it may be desirable

for security experts to be involved in negotiation, that might not always be the case. This lack of involvement however can affect the negotiations process negatively.

The findings presented and analyzed in the previous paragraphs all help to support the reality of the problems faced as companies transition between BOF4WSS phases (or any general cross-enterprise negotiations task really). Mark's statement in response to the question about transition problems sums it up aptly as he expressed, "Oh, I've seen that, and you're exactly right, that is the way it happens, it takes months, possibly years in some circumstances". This quote captures the seriousness of the transition problems highlighted in this research.

2) Risks drive security actions and requirements: To ease difficulties in the initial matching and comparison of security actions and requirements across enterprises, the Solution Model proposed the use of a shared risks catalogue. A common risks base would be key to allowing for automated matching using a tool. Central to this proposal was the idea that risks are the core drivers for security actions. This notion was supported by literature surveyed in [12] and thus embodied in the resulting ontology. With appreciation of the importance of this notion to the Model and resulting software tool (that is, SASaCS), it was chosen for assessment in the interviews.

Reporting on the data gathered, a majority of professionals supported the 'risk-driven' notion. Feedback ranged from, "it always stems from risks and understanding risks, risk management, risk evaluation, it really drives everything to be honest", to "driving security, a risk-based approach something I firmly believe in". Cost factors were also mentioned by one interviewee but these still related to underlying risks and their mitigation cost/benefit savings. Interviewee feedback therefore can be seen to give support to findings in our previous work in [12].

While accepting the role of risks as a driver for security, one interviewee expressed that a number of companies do not actually operate on a risk basis. Unfortunately, no examples were given as to what companies might do instead to define their actions. This reality is nonetheless a thought-provoking one in terms of the Solution Model because even though it is not ideal (interviewees and research from [12] point to a risk-based approach being best), if it is widespread, it might limit the adoption of the Model and tool.

The last important finding related to the communications benefit likely to result in using risks as a base for security-related discussions. Interviewee feedback highlighted that in using a risks base, security professionals and business persons (involved in negotiations) alike could understand what was at stake (impact to organization and so on). From this research's perspective, this is beneficial for two reasons. Firstly, if business-level personnel do engage in security negotiations (as alluded to in the theme above), using a language they will understand would give them the necessary insight into the process. And secondly, business persons are typically the budget holders (John and Mark emphasize this) therefore again, they have to understand the need for security for funds to be released to implement security actions.

3) Likelihood of sharing detailed information on risks and risks' treatments: The Solution Model and BOF4WSS requires that business partners share a great amount of information on common risks faced, factors (including, laws, organizational policies, and so on) that influence/motivate security actions, and security actions themselves (whether they are geared towards risk mitigation or otherwise). With appreciation of the possible inherent difficulties accompanying this task (such as companies not wanting to share such information), this evaluation theme focuses on how realistic is it an expectation.

The conclusions from the data analysis in this segment were less clear, and even in cases where professionals felt that information sharing was realistic, they still placed a number of conditions on sharing. For example, some stated that once the data requested was at a relatively high level and did not go into specific vulnerabilities or impacts to the organization, it would be feasible. This was an intriguing finding because the structure of the risks catalogue and data in SASaCS does to some extent ask companies to define specific vulnerabilities that constitute a risk. This might therefore require the catalogue structure to be modified slightly to show less detail, or finding scenarios where parties were likely to be open and the structure could be accepted as is.

Supporting the opposite view, the feedback did observe that in some situations, companies might refuse to give much information to partners and cite confidentiality reasons. Overall however they were two prerequisites identified that would increase likelihood of information sharing. These were, trust and an existing relationship between companies. Mark states, "a lot of companies, particularly in private sector are unlikely to do that unless you've got that trust". This shows a significance of existing relationships and trust to the Solution Model, similar to that necessary for the framework.

4) The ultimate use of the Model and tool: The SASaCS tool is a software implementation of the Solution Model. As such, it aims to streamline a number of tedious, repetitive and long-winded tasks, and thus, significantly ease transition between framework phases. The evaluation of the Model, largely by way of the tool, was therefore imperative in these interviews. To conduct this evaluation, the tool prototype was demonstrated to interviewees and then questions were asked. Below the feedback and analysis results are presented.

In response to questions regarding the tool's usefulness in supporting phase transition, interviewees felt that it was a very useful approach and system. John stated, "I think it would be really useful. Having seen it, I think the penny

has dropped for me, I think this could be very powerful, very useful. I think this would help a lot". Furthermore he expressed, "And it would accelerate the adoption of technology solutions and this framework". John made this statement because he felt that in business today, collaborations are somewhat technology-focused and what inhibits projects is the discussion and agreement difficulties arising from the business and legal sides. The tool to him, was seen to help these sides by considering security at a higher level, communicable to people at this layer (business or legal professionals for example).

Mark was another professional who strongly supported the tool's usefulness. He commented, "a tool that helps bring that [core negotiation aspects] directly onto the table, it makes that time together far more productive". Such opinions as those mentioned here and above give evidence to support the increased productivity achievable by using the tool (and the underlying Solution Model proposed). Matthew reinforces these point as he states, "I can think of projects that it probably would have shaved off months, in terms of the initial stages of that project, had they thought to do this earlier on".

When questioned about whether they (interviewees) would use the tool in such a negotiations scenario, a majority of subjects said that they would consider it—increased productivity being cited as the prime factor. Proponents also stated that the novel benefit with the Model and tool was that they laid out companies' security positions in a clear and direct format, and forced them to agree or disagree on positions/postures. Regarding the automated identification of conflicting security actions for risks, John stated, "you almost know straight away that the collaboration is not going to work unless someone changes their posture or they agree to something". The tool can therefore save time for companies in this regard (a feasibility level) also.

From a usability perspective, generally positive feedback was recorded. Perceived benefits related to good accessibility due to the use of a browser-based report format, and the ease at which security actions from companies could be compared. Shortcomings mentioned included the need for increased flexibility in tool output (such as, additional buttons and more options on screen). These are accepted as areas for improvement in moving from a prototype to construct a full version of SASaCS.

Even though interviewees affirmed the tool's usefulness in significantly supporting the phase transition, some noteworthy shortcomings were identified. Critiquing on the higher level data present in the tool, Luke states, "it seems useful with the caveat that it might hide stuff away from the decision makers". To remedy this, he suggests a drill-down functionality to allow more detail to be seen on treatments or risks. This feature would be used by security professionals involved in negotiations, whereas business-oriented decision makers might be happy with the current higher level information. Speaking objectively, this is a useful suggestion but if implemented it would have to be optional. This is because, as was identified in the previous discussion theme, all companies might not be willing to share detailed information. Trust, to some extent, again becomes a factor.

Another observation mentioned was the dependence of the tool on the quality of the input data. "It is the input data's quality that is going to impact on the influence [of the tool]", Luke stresses. Matthew also supported this fact. To reply to this point, we accept it as an issue, however little can be done beyond giving guides and on screen tooltips to companies and users. It is assumed that companies would appreciate the productivity benefits when quality data is provided, and therefore use the Model and tool as suggested. Inadequate provision of information by some partners in a collaboration might even act as an indicator to other companies as to how serious partners are regarding collaboration and collaboration security.

5) Summarizing Solution Model analysis: In the following paragraphs, the findings presented and analyzed above are summarized in a *theme-by-theme* fashion. The conclusions drawn are then used to determine the degree of support for the *area* highlighted at the beginning of Section IV-C.

The first theme of analysis related to determining the severity of the transition problems that motivated the Solution Model's design. From the data, it was clear that a majority of professionals appreciated the problems (largely drawing on their own experiences), and viewed them as quite serious issues within projects. Additional issues were even highlighted relating to companies themselves not being clear on the exact motivation for security actions, and inexperienced personnel being involved in negotiations. Considering these points in light of the area under analysis, they can therefore be seen to support the seriousness of transition problems, especially relating to the great deal of time consumed, and lack of productivity.

The Solution Model operates on the premise that security risks drive security actions and security requirements. The validity of this premise therefore directly affects the viability of the Model and resulting system/tool. Based on the data, most professionals supported this premise and viewed it as the best way forward. Furthermore, it was seen to have additional uses because the notion of a risk was viewed as a key communications tool that could give business persons the necessary insight into security. One contrary point to risks as a driver was that a number of companies actually do not operate on this basis. Without any clear indication of a standard, well-justified process to identify actions however, little could be done to address this issue. With respect to supporting the viability of the Solution Model therefore, the data was seen to strongly support a risks base to security actions.

For the Solution Model to work, companies are required to share detailed information on risks related to the scenario,

influential factors in risk treatment, and defined security actions. On assessing the likelihood of that occurring, the analysis conclusions were not clear. Some professionals regarded it as realistic, whilst others did not. Possibly the most noteworthy finding here however was that trust and an existing relationship were cited as factors that might increase the likelihood of this information being shared. This is an acceptable prerequisite as it largely fits in with the updated target scenarios of BOF4WSS outlined at the end of Section IV-B. Assuming an atmosphere with trust and an existing relationship therefore, the interview findings can be seen to support an enhanced level of information sharing, and thus to some extent, the viability of the Model.

In investigating the Solution Model by way of the tool, the most significant question would have to be centred around the ultimate strength of the process and tool itself. In response to this question, professionals gave very positive feedback and affirmed the usefulness of the tool in significantly easing cross-enterprise security negotiations. The Model and tool were especially seen to accelerate adoption of technology solutions, and increase productivity and reduce time spent in negotiations. Furthermore, one professional saw it as beneficial to the overarching framework such that it would accelerate its adoption. This formed a critical point because it highlighted that research into support systems (such as the Solution Model and tool) could impact on the adoption of BOF4WSS.

Another important advantage is the fact that by requesting information on motivational/influential factors *before* companies meet, entities will have to find clear justifications to support their security actions. This directly helps to address the issues related to incomplete information and weakly justified motivational factors identified in the transition problems theme. Reflecting on the analysis area therefore, the findings and conclusions from this theme strongly support the viability of the Model and tool in supporting phase transition. There might be some slight improvements that can be made (including, drill down functionality, modifying structure of risks data in the catalogue and SASaCS) but these were not seen to seriously affect the use of the tool or viability of the Model.

In summary, the findings gathered provided a solid degree of support for the viability of the Solution Model in greatly aiding the transition between Requirements Elicitation and Negotiation phases of BOF4WSS. Trust and existing relationships between parties also played an important role, however this is acceptable as it coincides with the updated target scenarios of the framework.

Lastly, as this section represents the second evaluation of the Solution Model and tool (the first was the compatibility assessment in [4]), the findings and conclusions of the two evaluations were compared for any points of interest. One important observation was found. This was based on the fact that constraints (laws, obligations, policies, and so on) were seen as an additional driver of security actions in [4], whereas in this evaluation security professionals only mentioned risks. Although this leads to no clear conclusion, because the Model and tool by nature should be comprehensive, they should arguably accommodate both cases. Critically speaking therefore, the viability of the Model and tool can be regarded as negatively affected because currently they only use a risks base (and thus will only automate handling of risk-based security actions). Possible ways that constraints could be included in automated handling were previously discussed in [4].

Even though the negative feedback mentioned above harms viability, the strong support for the risks base and the tool in general supplied by industry-based professionals was felt to outweigh this aspect. Future work towards automated handling of constraints will be pursued only to ensure that the Solution Model and tool are as comprehensive as possible. This would allow them to handle a greater number of situations in which they are required to support cross-enterprise negotiations.

V. Conclusion and Future Work

In this paper we reported on the results from an evaluation conducted on two of our previous research proposals; namely, BOF4WSS and the security negotiations Solution Model and Tool used to support it. Generally, findings were seen to support the framework and Model/tool as useful, viable and practical approaches in addressing the issues they target. There were however some limitations, particularly related to applicable scenarios for the framework, and contentions regarding security actions and their core driving factors. These were important but not viewed as factors that seriously undermined these research proposals.

The next step of this research is to build on the insights and favourable findings of the initial assessments, and conduct the final evaluation process. This evaluation would constitute a thorough case study analysis where real-world companies would be observed using BOF4WSS and its supporting tools. This study would complement preliminary evaluations and allow for a much more comprehensive analysis. Furthermore, it would enable for clear, well substantiated conclusions to be drawn from this research.

References

[1] J. R. Nurse and J. E. Sinclair, "A Solution Model and Tool for Supporting the Negotiation of Security Decisions in E-Business Collaborations," in *5th International Conference on Internet and Web Applications and Services (ICIW)*. IEEE Computer Society, 2010, pp. 13–18.

[2] ——, "BOF4WSS: A Business-Oriented Framework for Enhancing Web Services Security for e-Business," in *4th International Conference on Internet and Web Applications and Services (ICIW)*. IEEE Computer Society, 2009, pp. 286–291.

[3] ——, "Securing e-Businesses that use Web Services — A Guided Tour Through BOF4WSS," *International Journal On Advances in Internet Technology*, vol. 2, no. 4, pp. 253–276, 2009.

[4] ——, "Evaluating the Compatibility of a Tool to Support E-Businesses' Security Negotiations," in *The International Conference of Information Security and Internet Engineering (ICISIE), under World Congress on Engineering (WCE) 2010*, vol. 1. Newswood Limited, International Association of Engineers, 2010, pp. 438–443.

[5] B. Hartman, D. J. Flinn, K. Beznosov, and S. Kawamoto, *Mastering Web Services Security*. Indianapolis: Wiley, 2003.

[6] O. Demirörs, Ç. Gencel, and A. Tarhan, "Utilizing business process models for requirements elicitation," in *The 29th Conference on EUROMICRO*. IEEE, 2003, pp. 409–412.

[7] J. S. Tiller, *The Ethical Hack: A Framework for Business Value Penetration Testing*. Boca Raton, FL: Auerbach Publ., 2005.

[8] M. P. Papazoglou, *Web Services: Principles and Technology*. Harlow, Essex: Prentice Hall, 2007.

[9] C. Gutiérrez, E. Fernández-Medina, and M. Piattini, "PWSSec: Process for web services security," in *The IEEE International Conference on Web Services (ICWS'06)*, Chicago, IL, September 2006, pp. 213–222.

[10] W. D. Yu, D. Aravind, and P. Supthaweesuk, "Software vulnerability analysis for web services software systems," in *IEEE Symposium on Computers and Communications*. IEEE Computer Society, 2006, pp. 740–748.

[11] S. Dynes, L. M. Kolbe, and R. Schierholz, "Information security in the extended enterprise: A research agenda," in *AMCIS 2007 Proceedings*, 2007.

[12] J. R. Nurse and J. E. Sinclair, "Supporting the Comparison of Business-Level Security Requirements within Cross-Enterprise Service Development," in *Business Information Systems*, ser. Lecture Notes in Business Information Processing, W. Abramowicz, Ed. Heidelberg: Springer, 2009, vol. 21, pp. 61–72.

[13] S. S. Yau and Z. Chen, "A framework for specifying and managing security requirements in collaborative systems," in *Autonomic and Trusted Computing*, ser. Lecture Notes in Computer Science, L. T. Yang, H. Jin, J. Ma, and T. Ungerer, Eds. Heidelberg: Springer, 2006, vol. 4158, pp. 500–510.

[14] J. Roy, M. Barik, and C. Mazumdar, "ESRML: a markup language for enterprise security requirement specification," in *IEEE INDICON*, Kharagpur, 2004, pp. 509–512.

[15] M. D. Myers, *Qualitative research in business and management*. London: SAGE, 2009.

[16] B. L. Berg, *Qualitative research methods for the social sciences*, 5th ed. London: Pearson International Education, 2004.

[17] C. Teddlie and F. Yu, "Mixed methods sampling: A typology with examples," *Journal of Mixed Methods Research*, vol. 1, no. 1, pp. 77–100, 2007.

[18] M. Todd, E. Zibert, and T. Midwinter, "Security risk management in the BT HP alliance," *BT Technology Journal*, vol. 24, no. 4, pp. 47–52, 2006.

[19] W. H. Baker, G. E. Smith, and K. J. Watson, "Information security risk in the e-supply chain," in *E-Supply Chain Technologies and Management*, Q. Zhang, Ed. Hershey, PA: Idea Group Inc., 2007, pp. 142–161.

[20] A. Baldwin, Y. Beres, S. Shiu, and P. Kearney, "A model-based approach to trust, security and assurance," *BT Technology Journal*, vol. 24, no. 4, pp. 53–68, 2006.

[21] F. Goethals, J. Vandenbulcke, W. Lemahieu, and M. Snoeck, "Different types of business-to-business integration: Extended enterprise integration vs. market B2B integration," in *E-Business Innovation and Process Management*, I. Lee, Ed. Hershey, PA: CyberTech Publishing, 2007, pp. 1–17.

[22] A. Singhal, T. Winograd, and K. Scarfone, "Guide to secure web services (NIST Special Publication 800-95)," National Institute of Standards and Technology (NIST), Tech. Rep., 2007.

[23] F. Satoh, Y. Nakamura, N. K. Mukhi, M. Tatsubori, and K. Ono, "Methodology and tools for end-to-end SOA security configurations," in *IEEE Congress on Services - Part I*. IEEE Computer Society, 2008, pp. 307–314.

[24] M. Tatsubori, T. Imamura, and Y. Nakamura, "Best-practice patterns and tool support for configuring secure web services messaging," in *IEEE International Conference on Web Services*. Athens, Greece: IEEE Computer Society, 2004, pp. 244–251.

[25] C. Van Slyke and F. Bélanger, *E-Business Technologies: Supporting the Net-Enhanced Organization*. New York: Wiley, 2003.

[26] PricewaterhouseCoopers LLP and Infosecurity Europe, "Information Security Breaches Survey 2010: Executive Summary," 2010, http://www.pwc.co.uk/pdf/isbs_survey_2010_executive_summary.pdf (Accessed 7 May 2010).

[27] E. W. Davis and R. E. Spekman, *The Extended Enterprise: Gaining Competitive Advantage through Collaborative Supply Chains*. Upper Saddle River, NJ: FT Prentice Hall, 2004.

[28] B. S. Sahay, "Understanding trust in supply chain relationships," *Industrial Management & Data Systems*, vol. 103, no. 8, pp. 553–563, 2003.

www.iariajournals.org

International Journal On Advances in Intelligent Systems
ICAS, ACHI, ICCGI, UBICOMM, ADVCOMP, CENTRIC, GEOProcessing, SEMAPRO, BIOSYSCOM, BIOINFO, BIOTECHNO, FUTURE COMPUTING, SERVICE COMPUTATION, COGNITIVE, ADAPTIVE, CONTENT, PATTERNS, CLOUD COMPUTING, COMPUTATION TOOLS
issn: 1942-2679

International Journal On Advances in Internet Technology
ICDS, ICIW, CTRQ, UBICOMM, ICSNC, AFIN, INTERNET, AP2PS, EMERGING
issn: 1942-2652

International Journal On Advances in Life Sciences
eTELEMED, eKNOW, eL&mL, BIODIV, BIOENVIRONMENT, BIOGREEN, BIOSYSCOM, BIOINFO, BIOTECHNO
issn: 1942-2660

International Journal On Advances in Networks and Services
ICN, ICNS, ICIW, ICWMC, SENSORCOMM, MESH, CENTRIC, MMEDIA, SERVICE COMPUTATION
issn: 1942-2644

International Journal On Advances in Security
ICQNM, SECURWARE, MESH, DEPEND, INTERNET, CYBERLAWS
issn: 1942-2636

International Journal On Advances in Software
ICSEA, ICCGI, ADVCOMP, GEOProcessing, DBKDA, INTENSIVE, VALID, SIMUL, FUTURE COMPUTING, SERVICE COMPUTATION, COGNITIVE, ADAPTIVE, CONTENT, PATTERNS, CLOUD COMPUTING, COMPUTATION TOOLS
issn: 1942-2628

International Journal On Advances in Systems and Measurements
ICQNM, ICONS, ICIMP, SENSORCOMM, CENICS, VALID, SIMUL
issn: 1942-261x

International Journal On Advances in Telecommunications
AICT, ICDT, ICWMC, ICSNC, CTRQ, SPACOMM, MMEDIA
issn: 1942-2601