Data Model Development for Security Information Sharing in Smart Grids

Rafał Leszczyna, Michał R. Wróbel
Gdańsk University of Technology
Faculty of Manage ment and Economics Gdańsk, Poland

Abstract

The smart grid raises new security concerns which require novel solutions. It is commonly agreed that to protect the grid, the effective collaboration and information sharing between the relevant stakeholders is prerequisite. Developing a security information sharing platform for the smart grid is a new research direction which poses several challenges related to the highly distributed and heterogeneous character of the grid. In this paper an approach to developing a data model for security information sharing platform for the smart grid which addresses these questions is presented together with the results of its application.

1. Introduction

The smart grid is a new form of electricity network which enhances the 20th century power grid and takes advantage of Information and Communication Technologies (ICT) to enable two-way power and information flows and to create an automated and distributed advanced energy delivery network [1].

The traditional grid is based on centralised power generation and electromechanical solutions. Monitoring of the grid as well as its potential restoration in case of a failure is performed manually which gives the operators only a limited control. In the smart grid, on the other hand, power generation centres are distributed and interconnected with a power and communication network which utilises digital solutions and sensors. This results in smart grid capabilities of self-monitoring and self-healing, as well as its high adaptiveness [1].

The anticipated benefits of the smart grid include [1]:

- Improved power reliability and quality
- Self-healing and increased resilience to disruption
- Predictive maintenance
- Facilitated deployment of renewable energy sources and distributed power sources
- Automated maintenance and operation
- Increased consumer choice

However there are also new challenges related to the development of the new domain. One of them is to address the security and privacy concerns which are the consequence of the high dependence of the grid on the ICT and its interconnection with the Internet. Each network connection of the grid constitutes a potential entry for a cyber-attack and every network layer and the technology used may become its possible target. Moreover, because the smart grid is a complex system of distributed and interconnected systems, it presents an exceptionally large attack surface [2].

The new grid is exposed to the whole myriad of cyber-threats which, even worse, evolve very quickly. Botnets, zero-days, Distributed Denial of Service Attacks (DDoS) or Advanced Persistent Threats (APT), are only few examples of threats which emerged or advanced significantly in the last years. There are also completely new threats inherent to the smart grid domain. These, for instance, include the attacks on the smart grid metering infrastructure. Compromising a smart meter opens a way for reaching other smart grid devices, such as smart thermostats, appliances, charging stations, because they are all interlinked. Furthermore, the localisation of some of the smart grid components in public places or at the end user's facilities exposes them to a nearly 24/7 potential attacker activity.

Deploying the standard and established security solutions such as firewalls, Intrusion Detection/Prevention Systems (IDS/IPS) or anti-malware tools is prerequisite for securing the grid. However, to counter the evolved and highly sophisticated threats such as the APT or DDoS, advanced technologies are required. It is commonly agreed that in achieving information security objectives effective collaboration between the relevant stakeholders (governments, energy providers, customer organizations etc.) is paramount. The use of partnerships and information sharing has become critical to the smart grid security as among the others it enables faster and more accurate response to new threats, providing measurable cost savings [3].

Project DEnSeK (Distributed Energy Security Knowledge) [4] addresses this situation by providing a platform for the security knowledge exchange between companies of the European energy sector. The project aims at improving the security and resilience of the new energy infrastructure against cyber-threats by [4]:

- Establishing a European Energy ISAC (Information Sharing and Analysis Centre) which will enable interactive and real-time knowledge and information sharing between all involved parties
- Deploying a Situation Awareness Network to monitor threats within European energy networks
- Developing an Information Sharing Platform which will form a trusted network to liaise between the ISAC and its members

The crucial part of the development of the information sharing platform is the proper identification and representation of the data exchanged in the platform. Besides the standard problem of assuring completeness and consistency of the data, there are also challenges specific to the smart grid security domain. It is well-known that user involvement is a key factor in the data identification and modelling process. This is particularly important when various users have some common and some unique data needs.

The smart grid is characterised by large geographical distribution and the involved enterprises or institutions are often located in geographically distanced areas (e.g. in the countries on the opposite borders of a continent). Moreover the companies and institutions are very heterogeneous in regard to their size (from small to very big) and business (i.e. energy generation, distribution, utilities, vendors, research and academia, government, international agencies, standard development organisations) and often represent different interests. As a result it is very difficult to establish good communication with all the representatives which is necessary in obtaining their input and feedback.

Another question regards the fact that in the information security domain a part of the exchanged information will be machine-generated. For instance the automatically created reports of IDS/IPS or anti-malware tools will be sent between the interested parties.

Finally, the challenge which was faced during the development of the model regarded the representation of the data. Over sixty years of data modelling domain resulted in a very impressive set of methods, models, languages, representations and tools. In this context choosing the right for the domain of smart grid security becomes non-trivial and requires to be preceded with a thorough analysis.

In the paper an approach to developing a data model for security information sharing platform for the smart grid which addresses these questions is presented (Section 4). Sections 2 and 3 show the results of the analysis of the data modelling domain. In Section 2 the most important concepts and classifications of the data modelling are shortly described while Section 3 comprises an overview of related works. The paper concludes with a presentation of the data model developed by applying the presented approach.

2. Data Models

2.1. Definitions

Navathe defines a data model as "a set of concepts that can be used to describe the structure of and operations on a database" [5]. In this definition the term is referred to a database as it originally comes from the field of database design where it was introduced in the 1950s. The definition of Navathe is well established in the field and commonly agreed.

Software engineering understands the concept of the data model as a visual representation of the objects of interest to a business and the relationships between them.

In this paper an amalgam of the above definitions is used in reference to a data model which is referred as a set of concepts that can be used to represent the objects of interest to a business, the structure of the relationships between them and the possible operations on the objects. It needs to be noted that the concepts used in the representation include visual elements.

In the problem area described in this paper the main objects of interest are data exchanged during security information sharing.

Data models represent the following types of characteristics of the represented contexts [6]:

- Static properties – objects, object properties (sometimes called attributes), relationships amongst objects
- Dynamic properties – operations on objects, operation properties, and relationships amongst operations
- Integrity rules over objects and operations

Abstraction is an important concept in data modelling (and modelling – in general). It refers to "an intuitive technique which requires a modeller to define the essential features of a real world domain and represent these features in a different form" [7].

2.2. Taxonomy of Data Models

Data modelling has evolved through a number of generations, each of them reflecting an increased level of abstraction. As a result the following basic classes of data models can be distinguished [8]:

- Hierarchical data models
- Network data models
- Relational data models

- Semantic data models

The *hierarchical*, *network*, and *relational data models* are often referred to as the *classical data models*. The first two models are highly machine-oriented. In fact, the *hierarchical model* emerged from sequential access methods associated with the medium used for data storage at that time i.e. the magnetic tape. In hierarchical models data are organised into record types with fixed and ordered hierarchical links, which altogether constitute a tree that refers to the hierarchical organisation of one-to-many relationships [8].

The challenge encountered by the designers of hierarchical models was the representation of many-to-many relationships and the redundancy associated with multiple hierarchies. To answer this challenge the *network data model* was developed. Network models support complex structures and many-to-many relationships, by allowing for sets intersections [9].

Both, the hierarchical and network models provide only basic operations. The internal organisation of data needs to be maintained by users. To assure this, data independence is required, i.e. the separation of logical and physical (implementation) properties [8].

Relational data models are far more user-oriented. In fact the emergence of the relational data models is perceived a significant step forward from the previous data models. The basic data structuring tool used in the relational model is the *relation*, or *table*. A table specifies the structure, i.e. the attributes of a collection of so called *tuples*. Each tuple links the attributes with values. In contrary to the hierarchical and network models, the ordering of tuples in a relation is not important to the user of a database. Additionally, each table is independent, meaning that there are no explicit links between tables to represent relationships. The relationships are represented implicitly by matching attribute values in two tables [8].

Semantic data models are the results of further development in the field of data modelling which aimed at providing more user-oriented modelling flexibility and creating models which are independent from machine-oriented concepts such as pointers in the network model. Semantic data models focus on mechanisms to capture more domain knowledge in the structure of data. They are the response to the lack of expressiveness of relational models as far as semantics of the data are concerned. To extend the expressiveness and flexibility of models, semantic modelling introduced abstraction constructs, such as generalization and aggregation [9].

Entity-relationship (ER) *model* is the most widely known and acknowledged semantic data model. The fundamental modelling construct in this model is the notion of *entity type*, which reflects the common characteristics of a set of real world entities. The common characteristics of entities belonging to a given entity type are characterised by the attributes specified for an entity type, in the sense that all entities of a given type share all the attributes defining that type. The relationships are associations between different entity types [6].

As far as the main constructs used in the model i.e. the entities, attributes and relationships, are concerned, the model is highly conceptual, because they are completely independent from any implementation. As a result any ER model is easily transformable into a hierarchical, network, or relational model (although some semantics may be lost in the conversion). However, without additional information (which regards the semantics of data), it is not possible to transform any of the classical models to a semantically equal ER model [8].

2.3. Data Modeling Process

The main purpose of defining a data model is to facilitate the communication between business and technical personnel of a project in order to create a common understanding of a given problem. Based on the requirements gathered from the interested parties, data model delivers a representation of the static and dynamic properties of business processes [6]. In addition, data modelling provides high-level technical specifications of the project by determining the structure of the information stored in the information system.

Data modelling is performed in the early stage of a software development process. It has a significant impact on all the next phases of the development of a solution. The quality of a data model will affect both the value of the solution and the cost of the development process. Therefore it is very important to precede the design of a data model with a very thorough analysis of the problem domain.

User involvement is a prerequisite for the development of a high-quality data model. This is particularly important when various users have some common and some unique data needs.

In the most common approach to data modelling the following four phases are distinguished [5]:

- Business requirements analysis
- Conceptual data modelling
- Logical data modelling
- Physical design

Business requirements analysis is the process of identifying inconsistencies and deficiencies in the set of requirements gathered during the preceding requirements elicitation phase of software engineering process. Examples of such inconsistencies and deficiencies include missing

requirements, requirements conflicts, ambiguous requirements or overlapping requirements and unrealistic requirements. These shortages need to be removed in order to establish an agreed set of requirements which are complete and consistent. The final set of requirements must be unambiguous to enable further steps of system development. The analysis also gives the designers an insight into a situation and allows them to make decisions regarding the further steps of the development [10].

Although requirements analysis is acknowledged as a critical success factor of information system development for organizations, mistakes are frequent at the requirements stage. Two of these mistakes are the lack of understanding of the business by requirements engineers and the miscommunication between business people and systems analysts. As a result of these problems, information systems may not fulfill organisational needs.

Conceptual data modeling aims at providing a clear and unambiguous understanding of the analysed domain or problem. For this reason, conceptual models should be abstract and simple representations of the real world, which provide effective means for communication. Conceptual models are the first representations of the domain which describe the most important objects and relationships between them. These representations are independent from a particular implementation. Conceptual modeling involves abstraction and simplification of reality which results in that less important objects and relationships are omitted [7].

Logical data modeling results in obtaining the models which represent the organisation of data in the analysed domain that (in contrary to conceptual modeling) is dependent on a particular implementation. The modeling constructs used in logical data modeling are straightforward for designers and avoid physical details of implementation, but usually can result in a direct computer implementation [5].

In the Physical design stage, data modeling is not used. Physical design is based on making decisions in regard to direct processing of physical data. In database design these decisions concern the storage of data in terms of clustering, partitioning, indexing, or providing additional access or directory structures [5].

3. Related Work

There are numerous models, languages, representations and tools used in data modelling.

Bachman diagrams were the first representations proposed in the domain. In Bachman diagrams boxes are used to represent the entities while arrows reflect the relationships between them. The diagrams are straightforward in designing and provide a useful tool for visual modelling, but they are not implementation-independent as they are devoted to the database schema for systems based on the CODASYL data model [11].

In the 1970s, Natural Information Analysis Method (NIAM) was proposed by Nijssen who adapted the Falkenberg's framework based on n-ary and nested relationships, and roles depicted with arrowed lines. Nijssen enhanced the framework with the circle-box notation for objects and roles. This circle-box notation became later a standard which is used till now. NIAM presents high capabilities for describing business rules and constraints and often takes support of natural language in making the models comprehensible. NIAM introduces the notion of facts – the combinations of entities, attributes, domains, and relationships, which need to be described together [12].

NIAM contributed to a broader approach of conceptual modelling called Object Role Modeling (ORM). ORM uses objects (entities or values) to represent the modelled domain. Objects can play various roles (parts in relationships). The ORM notation in contrary to NIAM, does not use attributes, which results in the increased stability of ORM models and queries as well as its higher uniformity. ORM models can also represent more complex business rules. ORM introduces the notion of universe of discourse (UoD) i.e. the modelled domain. The knowledge about the UoD can be verified after the translation of ORM models into pseudo natural language which is possible in ORM with the use of mix-fix predicates [12].

There are also models based on binary semantic networks, where all relationships are constructed from elementary binary relationships. Binary semantic models include for instance the Semantic Binary Relationship Model (SBRM) which aims at supporting the efficient modelling of practical enterprises by combining “an organisationally simple basis (i.e. binary relationships) with the capabilities of semantic networks and logical integrity and deduction rules” [13]. Alternative examples of binary semantic models include the models of Abrial, Bracci et al. FORAL [14]. Other established high-level data models are for instance the extended relational model RM/T, the functional model DAPLEX, the semantic relational model, or SDM or the IAM approach [14].

The models described above constitute a basis of data modelling on which the majority of contemporary practical modelling efforts in enterprises are based. For instance it is widely known that the Entity-Relationship Diagrams, UML class diagrams combined with the Object Constraint Language (OCL), the Object Description Language (ODL) specified in the ODMG standard, or data description language in SQL, are used in the daily practice [14].

From more modern approaches the study of Berman and Semwayo desires attention [15]. The authors introduce a new modelling primitive, called a niche, which reflects the environment where "entities interact for a specific purpose, playing specific roles, and according to the norms and constraints of that environment".

Another new contribution to the data modelling domain is active modelling. While conceptual modelling describes the static properties on the real world, active modelling also includes its activities and changes over time. Active modelling delivers methods for the conceptualisation of constantly changing reality. Its purpose is to enable continuous learning about the world through the acquisition and integration of human knowledge. Active models are able to represent many points of view of different stakeholders [16].

Group Data Modeler (GDM) is a collaborative data modelling process proposed by researchers from the University of Arizona. It allows multiple participants simultaneously work on the model content at the level of entities, attributes, and meta-data. Each stakeholder can provide their own entity definitions, list of attributes and properties. The process is divided into five stages, including initial preparation, defining entities, attributes, meta-data and relationships. Research conducted by the creators of the method showed that the reconciliation of products at each stage led to a better understanding of the domain and the various points by all participants. GDM also assumes anonymity which is believed to be a trigger for more open reporting of critical comments.

Extreme Conceptual Modeling (XCM) is the process of gathering user requirements by developing a conceptual model in the agile manner. XCM utilises a number of techniques from various areas of software engineering, such as Function refinement tree, Use case model or Function decomposition table. It is divided into the three stages: requirements engineering, conceptual modelling and software representation. This method provides also the Execution model, which allows for transforming conceptual models into their software representations [17].

The approach presented in this paper is unique in that it is tailored to modelling of the data shared in a smart grid security information sharing platform, a part of which being machine-generated, and provides a solution to the communication problem with highly dispersed and heterogeneous stakeholders which hinders obtaining their input and feedback.

4. Data Model Development Approach

4.1. Design Challenges

There are several challenges which concern designing a data model for a smart grid security information sharing.

The major problem is related to the heterogeneity and a large geographical distribution of the grid stakeholders. The stakeholders i.e. energy generation, distribution, utilities, vendors, research and academia, government, international agencies, standard development organisations represent different business sectors, business models, interests (sometimes opposite) and forms of activity which results in a diversified level of involvement in new initiatives. Moreover these enterprises and institutions are placed in various geographical locations, some of them very distanced, which makes it difficult to organise a physical meeting during which all the stakeholders' representatives would be present.

In result it is very difficult to establish good communication with all the representatives which is necessary in obtaining their input and feedback regarding which information they would like to exchange and how should it be represented. This situation is even worsened by the sensitive nature of the shared data that makes the stakeholders (especially the enterprises) cautious and reluctant in providing the input and feedback. The trust between the companies has a great influence on the intentions and attitudes to security knowledge sharing.

Another question regards the fact that in the information security domain a part of the exchanged information will be reports automatically generated by security solutions such as IDS/IPS or anti-malware tools. This means that the developed data model needs to be flexible enough to accommodate the machine-generated contents.

Also the vast availability of methods, models, languages, notations and tools for representing the data model – the result of six decades of the data modelling field – at the end poses a challenge to the designers who describe an emergent domain such as the smart grid security. Deciding on the right representation becomes a non-trivial choice, which requires a thorough analysis.

4.2. Addressing the Challenges – The New Approach

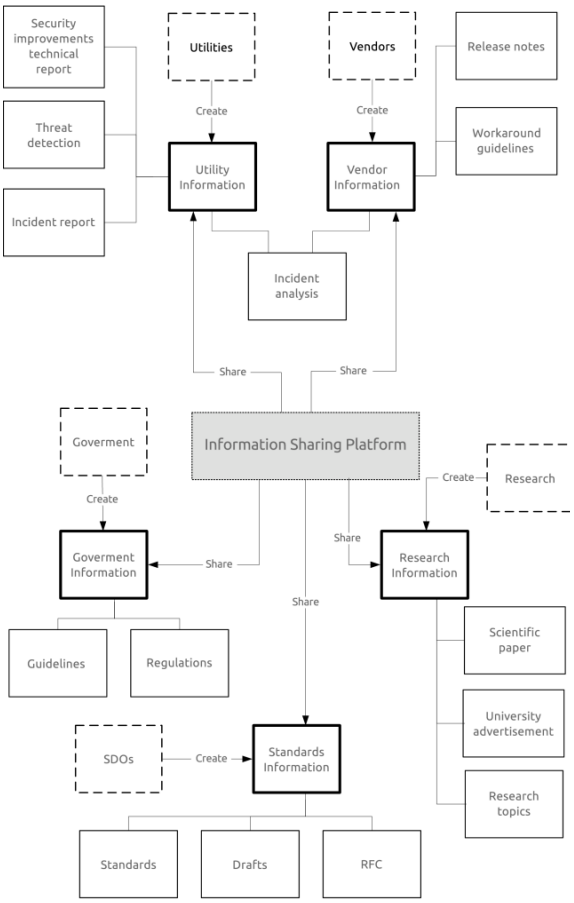

Figure 2. Very High-Level Data Model

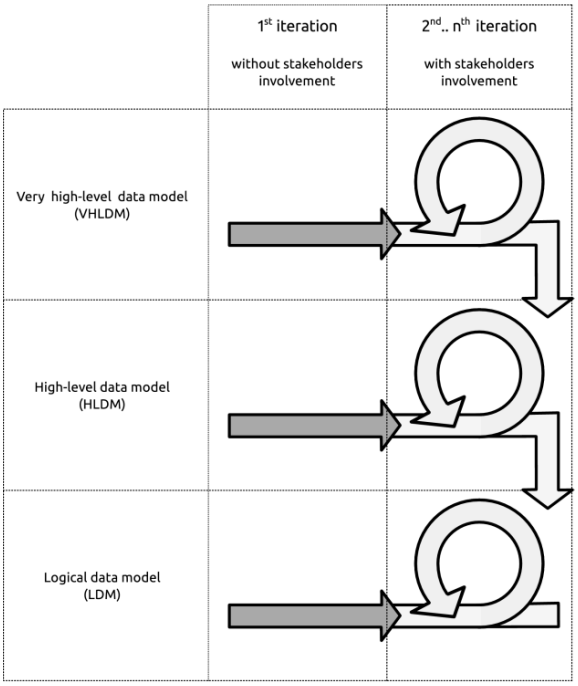

Figure 1. Data Model Development Process

In order to get orientation in the myriad of available data modelling methods, models, languages, notations and tools, the analysis of the domain was performed.

An overview of its results is presented in Sections 2 and 3.

As a result the proposed approach is an amalgam of the classical data modelling methodology where the four design phases are traversed (business requirements analysis, conceptual data modelling, logical data modelling and physical design) and an adaptation of the well-established iterative and incremental development model [18].

The iterative and incremental approach in data model development allows for obtaining a common vision of the shared data, the rules which govern them, possible operations and scenarios of use of the data, the roles of participants, and the whole universe of discourse in the environments where stakeholders are large international corporations, from various business areas, often competing with each other. The common position is derived in "small steps" – the intermediate, deliverable versions of the developed model, each one being an enhancement and an improvement of the previous one. This allows the developer to take advantage of what was being learned during the development of earlier versions of the model.

In the proposed approach the fundamental increments correspond to the phases of the data model design and include:

- Very high-level data model (VHLDM)
- High-level data model (HLDM)
- Logical data model (LDM)

In the first increment, based on the preliminary specification of information to be exchanged in the platform obtained during the requirements elicitation phase, a Very high-level data model is created (VHLDM). Based on the VHLDM, the High-level data model is designed which represents applications, users, rules and specifications of processing information assets. Finally the Logical data model is obtained which illustrates the relations between the entities and all business-related characteristics in the form of entity relationship diagram.

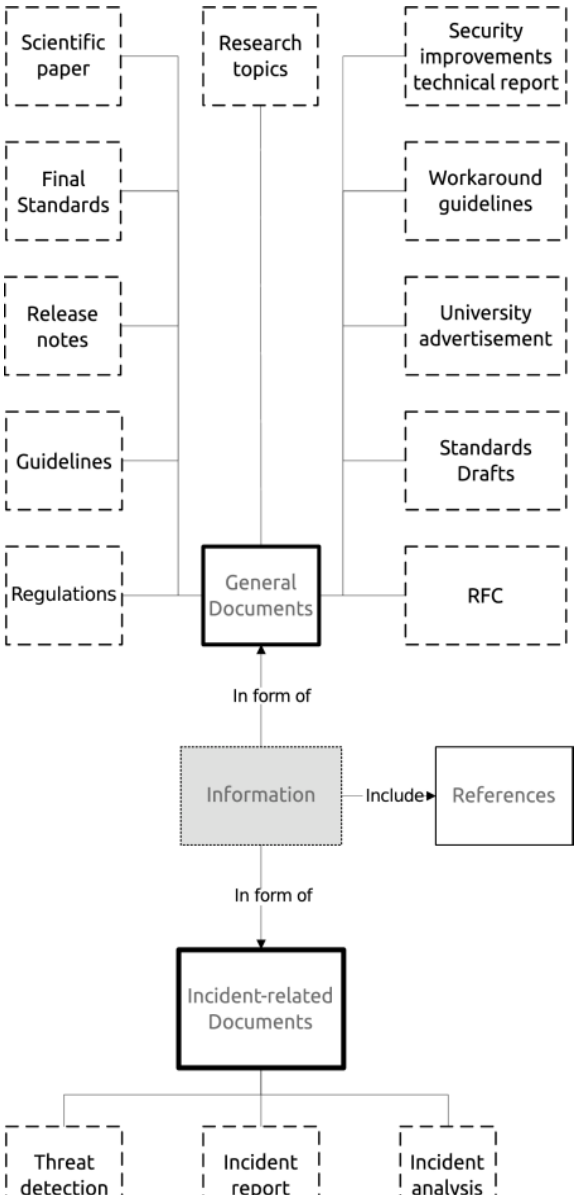

Figure 3 Groups of information exchanged in the security information sharing platform for the smart grid

For each increment at least three iterations need to be carried out. During the first iteration, data model is created without a direct involvement of stakeholder. It is based on the input received during a preceding increment and/or the analysis of available documentation, standards and other literature. Prepared according to the best knowledge and experience, the data model is presented to the stakeholders by means of electronic communication. This constitutes the second iteration. After obtaining the stakeholders' comments the third iteration is prepared where the document synthesising the input from the second iteration is submitted under a discussion during a physical meeting (a workshop). As a result, new models are created in subsequent iterations. The process can be repeated until the model is accepted by all stakeholders. Then the next increment is initiated. The data model development process is presented in Figure 1.

The final step in the development of the data modelling approach proposed in this paper was to address the question of model compatibility with the machine-generated contents (see the previous section). In this stage the available standard data representations for security information were analysed. The study resulted in the identification of the Intrusion Detection Message Exchange Format (IDMEF) and the Incident Object Description and Exchange Format (IODEF) as the direct representation of security related information, and the Dublin Core Metadata for general purpose documents. These representations were later integrated into the model (see Section 5) [19] [20] [21].

5. Data Models for a Security Information Sharing Platform for the Smart Grid

The approach described in Section 4 was applied to the development of data models for the security information sharing platform in the energy sector. This chapter presents the results of each increment.

5.1. Very High-Level Data Model

Based on the business model created in the preceding development stages, the main groups of the stakeholders have been identified. Further analysis resulted in the determination of types of information to be exchanged by the stakeholders. Each group of information is presented in Figure 2.

Utilities are companies directly involved in the production and/or distribution of the electricity. Their goal is to increase the security level by sharing knowledge about incidents and security issues. The data have been divided into four groups. Threat detection is sent immediately after the threat is detected. It contains brief information about the threat and should include at least the target of the attack. Incident report describes in details a detected security incident. It should contain information about the source of the attack, target, used tools, techniques and vulnerabilities. An incident report is sent at the time when the threat is under control. Incident analysis contains details of vulnerabilities and compromised security controls. It should include recommendations to improve security to avoid such incidents in the future. It is created based on the incident analysis. Security improvements technical report describes information about implemented security enhancements that could be undertaken by other shareholders.

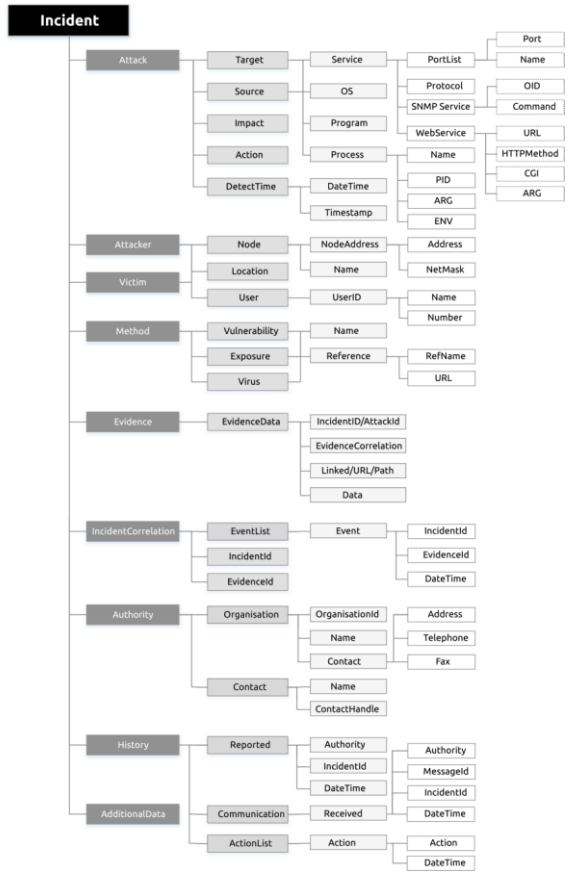

Figure 5. IODEF Data Model

The vendors group represents the manufacturers of hardware and software installed in the smart grid environment. Vendors are interested in sharing the release notes of new versions of tools and services used by at least one stakeholder. In this way they can encourage other stakeholders to acquire their products and services. Workaround description provides a guidance on how to solve problems with products and services provided by a vendor. It may be released in response to the threat detection or incident report sent by the utility. Finally, the incident analysis contains the analogous information to the one shared by the utilities.

Research and academia can share information on published or prepared scientific paper. The announcement should contain an abstract, keywords and a link to the full version. Other types of advertisements notify about scientific activities in the smart grid security domain, such as workshops, conferences or academic courses. Another type of information is research topic which informs about new research opportunities in the field of security of energy systems. It can notify about research teams, open positions, requests for MA or PhD theses, offers of cooperation, etc.

The government group includes all government agencies and institutions, both national and international, which can issue binding acts applicable in the field of smart grid. They can exchange guidelines and regulations for information security of power or industrial grids.

Standards developing organisations can share requests for comments, drafts and standards concerning new solutions and procedures, behaviours, research or innovations related to the security of smart grids.

5.2. High-Level Data Model

The analysis of VHLDM presented in the previous chapter allowed for grouping the information assets that may be shared in the information sharing platform.

The first group contains all information related to security incidents and threats. Extracting these data enables the automation in the processing of critical security information. The second group contains all remaining information which may be presented in the form of general document with additional metadata. The information groups are presented in Figure 3.

Based on the business model, channels of communication, rules, and participants were identified. These findings are encompassed in the High-level data model presented in Figure 4 (See the end of the paper).

The next step was to integrate into the HLDM the data representations which allowed for the compatibility with the machine-generated contents. As it was described in Section 4.2 the Intrusion Detection Message Exchange Format (IDMEF) and the Incident Object Description and Exchange Format (IODEF) as well as the Dublin Core Metadata were selected for this purpose. Since IODEF is fully compatible with IDMEF it is sufficient to integrate only this representation in the data model.

Dublin Core, standardised as ISO 15836:2009, includes the fifteen elements along with several dozen related properties and classes, which allow for describing in detail any document metadata [21]. Dublin Core is used for representing general documents exchanged within the security information sharing platform for the smart grid. The aim of the Incident Object Description and Exchange Format is to define common data format for describing and exchanging information about incidents between Computer Security Incident Response Teams (CSIRTs) [20]. IODEF is fully compatible with Intrusion Detection Message Exchange Format, which is orientated towards communication between Intrusion Detection Systems, yet extends it with objects enabling communication between people and teams. The IODEF Data Model is presented in Figure 5 [19].

These representations were integrated and tailored to the needs of the security information sharing

platform. Figure 6 presents the resulting detailed High-level data model.

5.3. Logical Data Model

The integration of the Dublin Core and IODEF at the level of the logical data model was based on the process proposed by Batini et al. [22]. First, models were analysed and compared in order to determine the relationships between concepts. Common properties and few conflicts were identified in this step. Then, after resolving conflicts, all concepts were placed on a common diagram. Such a model required restructuring, including a common convention for the names, adding the missing attributes and creating new relationships. Finally the new model was tested against the qualitative criteria: completeness, correctness, minimality and understandability. The LDM is presented in Figure 7.

6. Conclusions

The smart grid is a new application domain for information security which requires intense research efforts in various fields. One of the subjects which demand particular attention is the development of a security information sharing platform. This was identified as a priority area of the research and development in the grid.

In the paper a novel approach for proper identification and representation of the data exchanged in the platform was proposed that responds to the challenges inherent to the highly distributed and heterogeneous nature of the grid. The results of application of the approach (described in Section **Error! Reference source not found.**) proved its suitability and effectiveness.

The approach described in this paper can be applied to many application domains where similar challenges are faced. It is especially suitable in the contexts where communication problems arise due to the large geographical distribution of the involved enterprises and institutions as well as their high diversity in regard to size.

7. Acknowledgements

The study presented in this paper is based on work carried out in the DEnSeK (Distributed Energy Security Knowledge) project founded by the European Commission, Directorate-General for Home Affairs (Programme "Prevention, Preparedness and Consequence Management of Terrorism and other Security-related Risks" – CIPS, Project Reference: HOME/2012/CIPS/AG/ 4000003772) and partially supported from the project funds. The authors acknowledge the contributions of the DEnSeK consortium partners involving the development of the business model and their feedback on the data model developed by the authors.

8. References

[1] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart Grid — The New and Improved Power Grid: A Survey," IEEE Commun. Surv. Tutorials, vol. 14, no. 4, pp. 944–980, 2012.

[2] P. McDaniel and S. McLaughlin, "Security and Privacy Challenges in the Smart Grid," IEEE Secur. Priv. Mag., vol. 7, no. 3, pp. 75–77, May 2009.

[3] D. Feledi, S. Fenz, and L. Lechner, "Toward web-based information security knowledge sharing," Inf. Secur. Tech. Rep., vol. 17, no. 4, pp. 199–209, May 2013.

[4] "DEnSeK (Distributed Energy Security Knowledge) - project website." [Online]. Available: http://www.densek.eu/. [Accessed: 08-Apr-2014].

[5] S. B. Navathe, "Evolution Of Data for Databases Modeling," Commun. ACM, vol. 35, no. 9, pp. 112–123, 1992.

[6] Michael L. Brodie, "On the Development of Data Models," in On Conceptual Modelling, 1984, pp. 19–48.

[7] N. A. Karagöz, "A framework for developing conceptual models of the mission space for simulation systems," Middle East Technical University, 2008.

[8] J. Parsons, "Data Modeling," in Handbook on Data Management in Information Systems SE - 3, J. Błażewicz, W. Kubiak, T. Morzy, and M. Rusinkiewicz, Eds. Springer Berlin Heidelberg, 2003, pp. 49–77.

[9] W. D. Potter and R. P. Trueblood, "Traditional, semantic, and hypersemantic approaches to data modeling," Computer (Long. Beach. Calif)., vol. 21, no. 6, pp. 53–63, 1988.

[10] M. Chemuturi, Requirements Engineering and Management for Software Development Projects. New York, NY: Springer New York, 2013.

[11] C. W. Bachman, "Data structure diagrams," ACM SIGMIS Database, vol. 1, no. 2, pp. 4–10, Jul. 1969.

[12] T. Halpin, "Object-Role Modeling (ORM/NIAM)," in Handbook on Architectures of Information Systems SE - 4, P. Bernus, K. Mertins, and G. Schmidt, Eds. Springer Berlin Heidelberg, 2006, pp. 81–103.

[13] M. Azmoodeh, S. H. Lavington, and M. Standring, "The Semantic Binary Relationship Model of information," pp. 133–151, Jul. 1984.

[14] M. R. Kogalovsky and L. A. Kalinichenko, "Conceptual and ontological modeling in information systems," Program. Comput. Softw., vol. 35, no. 5, pp. 241–256, 2009.

[15] S. Berman and T. D. Semwayo, "A conceptual modeling methodology based on niches and granularity." 01-Jan-2007.

[16] P. Chen, B. Thalheim, and L. Wong, "Future directions of conceptual modeling," in Conceptual Modeling, 1999, pp. 287–301.

[17] E. Insfrán, V. Pelechano, and O. Pastor, "Conceptual Modeling in the eXtreme," Inf. Softw. Technol., vol. 44, pp. 659–669, 2002.

[18] V. R. Basili and C. Larman, "Iterative and Incremental Development: A Brief History," IEEE Comput. Soc., vol. 36, no. 6, pp. 47–56, 2003.

[19] H. Debar, D. Curry, and B. Feinstein, "RFC 4765 - The intrusion detection message exchange format (IDMEF)," Network Working Group, IETF. pp. 1–157, 2007.

[20] R. Danyliw, J. Meijer, and Y. Demchenko, "RFC 5070 - The Incident Object Description Exchange Format (IODEF)." 2007.

[21] Iso, "ISO 15836:2009 - Information and documentation - The Dublin Core metadata element set," Molecular Cell, vol. 36. pp. 885–893, 2009.

[22] C. Batini, M. Lenzerini, and S. Navathe, "A comparative analysis of methodologies for database schema integration," ACM Comput. Surv., vol. 18, no. 4, 1986.

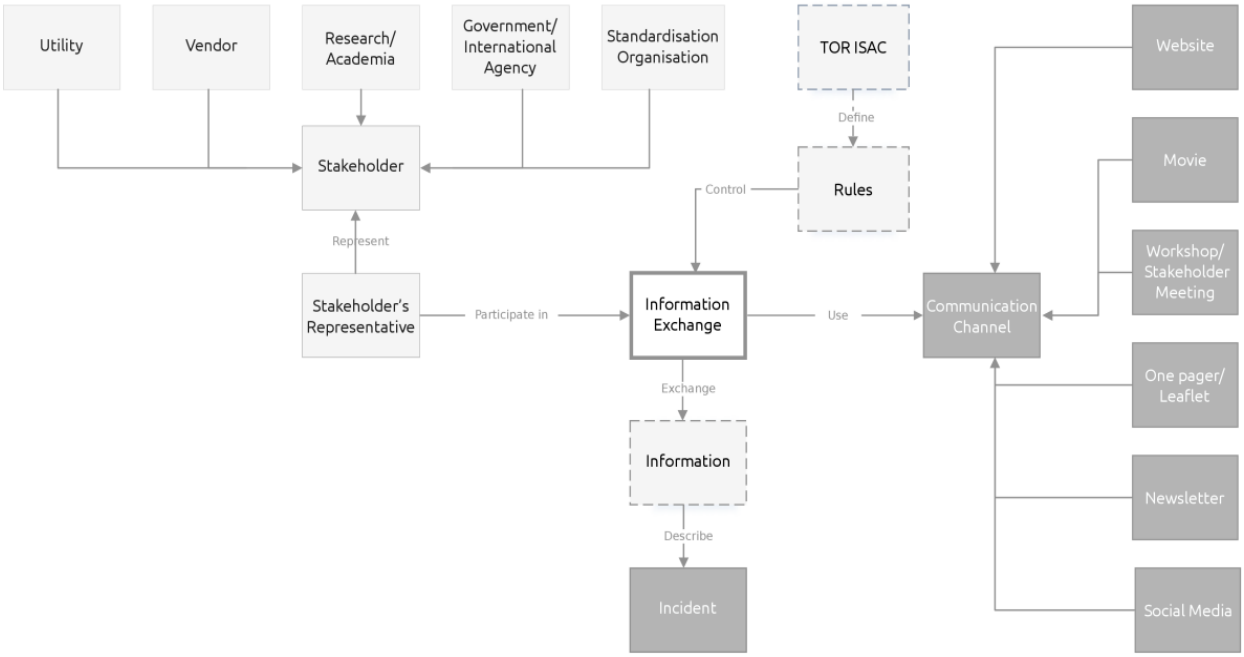

Figure 4. High-level data model

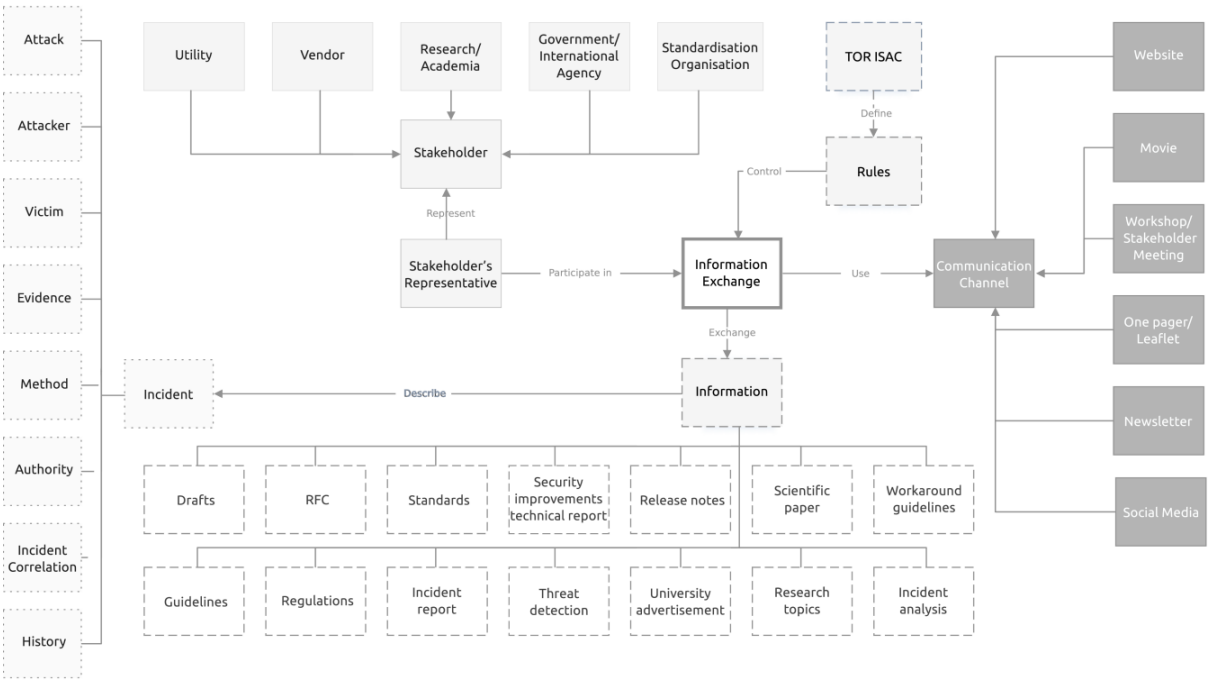

Figure 6. Detailed High-Level Data Model

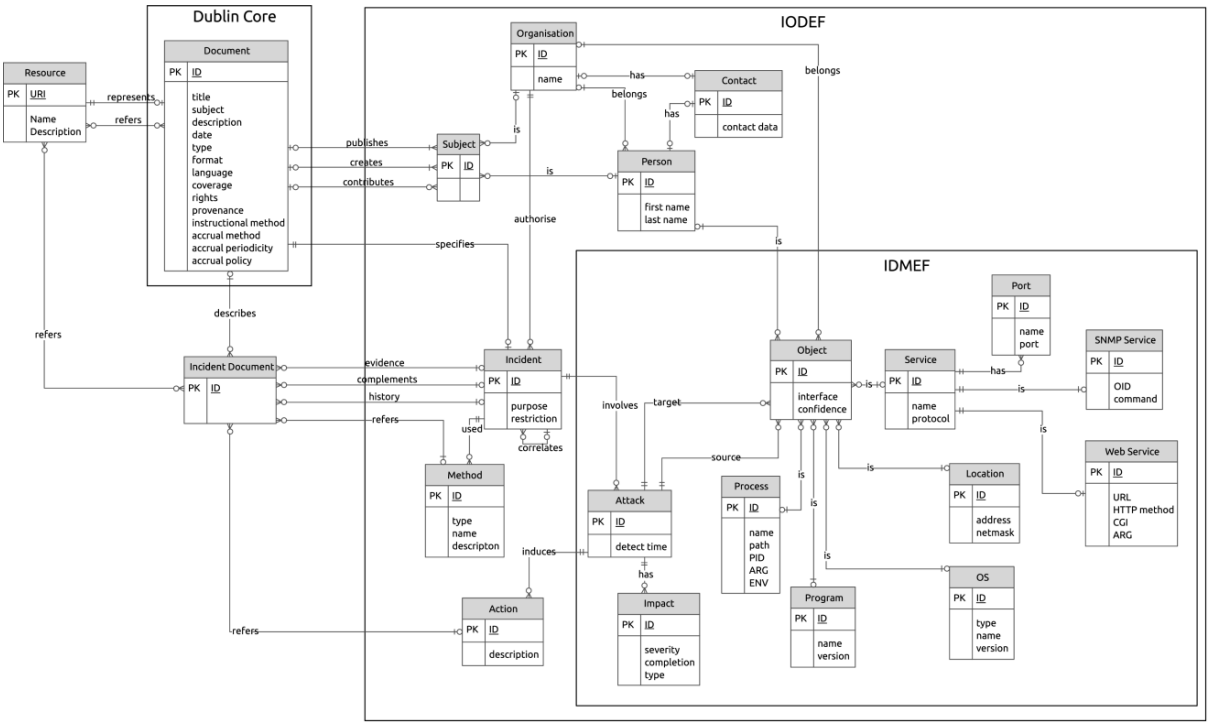

Figure 7. Logical Data Model