Threat Scenarios to evaluate Trustworthiness of Multi-agents in the Energy Data Management

Christine Rosinger[1], Mathias Uslar[1], Jürgen Sauer[2]

Abstract

Trustworthiness is a key aspect for coalition formation with agents in the energy domain. In this paper first results for developing a trust model for a multi-agent-based energy management system are given. First, the terms trust, reputation and trustworthiness are defined. After that an exemplary scenario of a trustworthy coalition formation of energy agents is described. The main result of this paper is the contribution of some threat scenarios which will be used for the evaluation of the trust model in a later stage.

1. Introduction

Improving the efficiency and reliability of the energy grid as well as using more renewable energy resources in the grid are goals that should be considered when implementing the future energy grid. Information and communication technologies (ICT) shall realize a fast information exchange of the actors – like consumers, producers or storages – of the smart grid. All these actors have to communicate with each other – some of them already do, but others have to be integrated for the first time. There are many actors which have a lot of interfaces and communication, which results in more threat potentials. Hence, existing security measures have to be re-considered and new security concepts have to be developed.

Another main aspect of the smart grid is distributed energy generation. Flexible, controllable loads, power plants and storages are connected to the grid and are integrated into energy data management (EDM) to achieve a higher automation ratio. One approach to realize automation is the concept of multi-agent systems. All actors of the EDM are represented as agents and communicate in the context of business processes. In this environment, uncooperative, malicious agents can emerge. As trust is a measure to monitor agents, the trustworthiness of the agents has to be considered to get a more robust system.

This paper provides an overview of some first results of developing a trust model in the project “Smart Nord”. After this introduction the terms trust, reputation and trustworthiness are defined in Section 2. Section 3 describes the project context where this trust model is applied to and contributes an exemplarily scenario of a trustworthy coalition formation. Nine threat scenarios for a multi-agent-based energy management system and a categorization for these are elaborated in Section 4. Additionally, evaluation questions and criteria are mentioned. The paper concludes with an overview and an outlook on further work.

2. Terms: Trust, reputation and trustworthiness

Trustworthiness does not consist of one temporal impression of an entity[3], but is a multi-faceted concept. Trustworthiness will be taken into account, if an entity has to rely on another entity to fulfill a task. The terms trust and security are often used with equal semantics in everyday language, but security is only one

[1] OFFIS – Institute for Information Technology, Escherweg 2, D-26121 Oldenburg,
email: christine.rosinger@offis.de, mathias.uslar@offis.de, Internet: http://www.offis.de.
[2] University of Oldenburg, Ammerländer Heerstr. 114-118, D-26129 Oldenburg,
email juergen.sauer@uni-oldenburg.de, Internet: www.uni-oldenburg.de
[3] The term entity is used in this paper to describe the concept of a human being or a system, like an agent.

of several facets of trustworthiness, as can be found in the following enumeration. Equally, the terms trust and reputation – which are part of the facet credibility – are applied alike. Trust is a subjective meaning of an entity; reputation is the common understanding of an entity, built by a group or a community (Jøsang et al 2007).

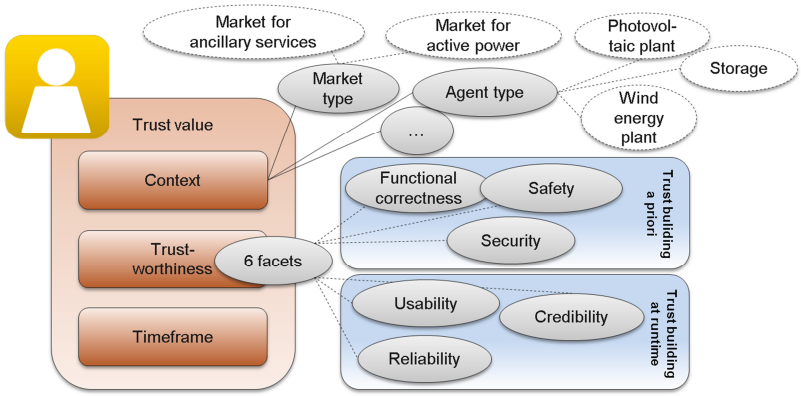

Figure 1: Abstract model of a trust value

Following (Steghöfer et al 2010), it can be distinguished between six facets of trustworthiness that can be adapted to the energy domain. Some of the different facets are more stable and can be determined a priori while others of them are completely dynamic and change during runtime. Some or all of these facets can be combined to one trust value/indicator. It is also possible to weight every facet differently. In Figure 1, a first draft of the modeled trust value of an agent is depicted. Besides these six following facets it is important to incorporate the context where the trust value occurred as well as the timeframe of occurrence (Rosinger et al 2012).

- **Functional correctness**: This facet describes the property of a system to be compliant with its functional specification. For the energy domain this means, e.g., if somebody executes the actions to power down a power plant it really shuts down. This can be verified by acceptance tests.
- **Safety**: This facet commonly defines that a system must not reach a state which harms itself or the environment. For the energy domain this means, e.g., that in an outage scenario of a power plant nothing must get damaged, neither something in the plant itself nor somebody who worked there. To check this, there are safety standards according to which systems can be certified.
- **Security**: The trust facet security determines that the security goals confidentiality, integrity, and availability must not be violated. In a trust model the security measures of the system, e.g., the agent of a wind power plant, have to be assessed. Security standards can be used for the certification of these systems.
- **Reliability**: This trust facet commonly means that despite occurring disturbances, failures, and prediction errors a system has to guarantee availability for a certain time. In the "Smart Nord" context, reliability means the probability with which a product is available within a product horizon under certain conditions (Blank/Lehnhoff 2013).
- **Credibility**: Generally, the facet credibility is the ability and the willingness of a cooperation partner to act beneficial, consistent, and transparent. In a multi-agent based system, credibility can be represented as the direct experience between two agents but also by the reputation. This facet will be described in more detail in Section 3.
- **Usability**: The facet usability assesses the offer of an efficient and effective user interface. At power plants ergonomic human-machine interfaces are used, however for the multi-agents scenario this facet does not play an important role.

3. Trust model: Example for trustworthy coalition formation

In this section an exemplary scenario for a trustworthy coalition formation with three unit agents[4] will be contributed as it will take place in the project Smart Nord (Sonnenschein et al 2012). The unit agents are actors of the multi-agent system and represent producers, consumers, and storages of power which form self-organizing coalitions to bid at the market. To achieve an optimal coalition formation several criteria – like economic or regulatory factors – will be taken into account. Another criterion is the trustworthiness of the agents which will support a trustworthy coalition formation. A trust model in order to deal with the trustworthiness of the agents will be developed in the project. The scenario is illustrated in Figure 2.

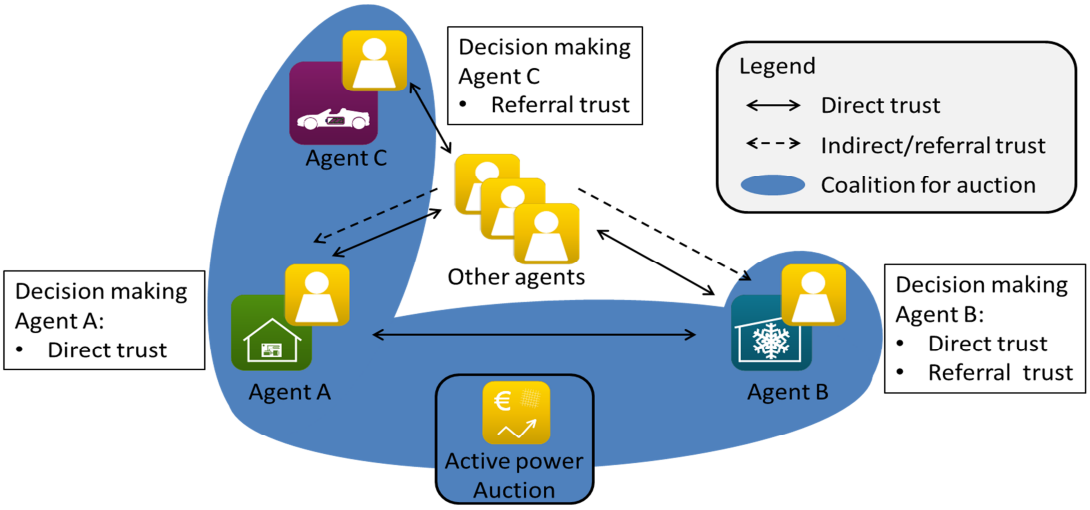

Figure 2: Exemplary scenario for trustworthy coalition formation

Agent A (a combined heat and power plant), agent B (a cooling house) and agent C (a battery storage electric car) want to bid in an auction for active power at an electricity market. Since the individual agents are too small to fulfill the full auction, every agent has to find several partners to form a coalition. In order to simplify the scenario it is assumed that the output of the three agents matches the proposed auction.

Agent A and B had lately a coalition with great success and some coalitions in the recent past, almost all proposed promises were satisfied by both agents in these past auctions. So they have a direct trust relationship. Thus, agent A decided to make a coalition with agent B without any other recommendation because these past coalitions satisfy his trust goal function.

Agent B is less trustful and decides to take additional recommendations (referral trust[5]) by different agents into account with which he and also agent A has direct trust relationships. With these referral trust values, Agent B forms for example a weighted mean value as it is, e.g., described in (Kiefhaber et all 2010). The trust value which agent B applies for the decision making consists of two parts: On the one hand, his own experience with agent A has a high weight. On the other hand, the indirect or referral trust of other agents will be weighted as agent B trusts the specific other agent.

Agent C has neither a direct trust relationship to agent A nor to agent B. Hence, the decision making process will only be supported by referral trust of agents that have a direct trust relationship with agent A and/or agent B. The values are also weighted by the trust in the referring agents.

[4] Unit agents represent different power plants and form coalitions to participate at a power market.

[5] The term referral trust is described in (Jøsang et al 2007).

This example refers only to the experience between two agents for past actions which is modeled by the trustworthiness facet credibility mentioned in Section 2. All other facets can also be taken into account for the trust building process between two agents. If, for example, the status of the implemented security measures is a big issue for an agent, the security trust value can be incorporated or get a greater weight. The trustworthiness consists (as described in Section 2) of six facets. All facets can, but do not necessarily have to, be involved. Additionally, the context[6] of the trust value and the timeframe when the experience of the referring agent occurred, have to be considered in the trust building process (See also Figure 1).

4. Evaluation: Using threat scenarios

In the project different threat scenarios have been identified that will be used in the evaluation process of the aforementioned project "Smart Nord" to assess the trust model. Firstly, this section gives a short overview of different motivations of attackers. After that, threat scenarios for the critical infrastructure energy and especially in the context of the project "Smart Nord" will be derived on the basis of the attackers' motivations. Finally, to evaluate the trust model, some evaluation criteria for the threat scenarios are described.

4.1 Motivations of attackers

For IT attacks, there is a common classification of different motivations of attackers which are described in the following bullet point list, mainly based on (Geschonneck 2008):

- **Social motivation:** Attackers with social motivation are like youth gangs where the acknowledgement of the mates is an important factor. Another social motivation is the resentment towards partners.
- **Technical motivation:** Attackers with technical motivation want to accelerate the security process by revealing vulnerabilities and gaps in systems or tools.
- **Political motivation:** Attackers with political motivation try to present their political believe while for example altering, defacing, or manipulating websites.
- **Financial motivation:** Attackers with financial motivation attempt to enrich themselves. This can be achieved by software piracy, financial fraud, or industrial espionage. In contrast to the first three attacker types these financial attackers would not boast about their actions.
- **Governmental political motivation:** Attackers with governmental-political motivation are instructed by the government to monitor and to get information or even alter information of other governments or major commercial enterprises.
- **Inadvertent motivation:** The previous mentioned motivations occur deliberately. Some threats or "attacks" arise out of inadvertent actions like carelessness, equipment failure, or natural disasters (IEC 2007).

4.2 Threat scenarios by multi-agents in the energy domain

As seen before, attackers have motivations and use security vulnerabilities to penetrate the system and to reach their goal(s). They are endangering the compliance to different security goals[7]. In the operational

[6] The context in this scenario is an auction for active power.

[7] The common security goals are confidentiality (keeping secrets), integrity (data without manipulation) and availability (durable running system) which are summarized and abbreviated CIA. Additionally the security goals authenticity (unforgeable identity) and non-repudiation (no withdrawal) which are both a part of integrity are considered.

scenario of the project are three different agent types –unit, market[8], and grid[9] agents – which have different goals as malicious agents. Another important factor in security is the attacker's origin, i.e. if he is an intrinsic or extrinsic attacker. In the following bullet point list several threat scenarios from the energy domain and multi-agent context are described. This threat identification is derived by common risk analyses of security standards like BSI[10] or NIST[11]. As presented in Table 1, these scenarios cover different categories in various combinations.

- **Threat scenario 1 (TS1): Identity fraud**
  A unit agent fakes his identity. He pretends to be a flexible storage but he is a weather-dependent photovoltaic power plant instead. He bids with his coalition at a market. They get the acceptance and the schedule is set. In the worst case – for him and the coalition – the photovoltaic power plant cannot fulfill the schedule. In the best case for him the unit can randomly accomplish the task and perhaps, additionally, undertake another task. In this case this malicious agent has an economic advantage because he gets more money as if it is due to him.
- **Threat scenario 2 (TS2): Unreliable agent**
  A unit agent sometimes fulfills his tasks in his coalitions and sometimes not. This threat is perhaps an inadvertent action. Another execution of this threat is that the unit agent leaves and resigns the coalition shortly before completing of the coalition formation. If this threat is not an inadvertent action there can be a financial reason.
- **Threat scenario 3 (TS3): Extrinsic attacker**
  An extrinsic attacker gets access to the multi-agent system by a security gap. He is, e.g., assigned by a foreign government/organization to acquire and perhaps alter information. Eavesdropping techniques (e.g., man-in-the-middle) can be used to get confidential information or manipulating techniques to alter data which can perhaps lead to damage in the grid. This can be a kind of cyber terrorism.
- **Threat scenario 4 (TS4): Political destruction**
  An extrinsic attacker hijacks an existing identity of a grid operator agent and tries to paralyze the grid, e.g., because of his political belief (for example, that he cannot accept the current energy roadmap). For this he distracts different components for example which perhaps force the electric grid to fail.
- **Threat scenario 5 (TS5): Organized crime**
  A market agent whose task normally is to communicate with coalition forming agents regarding products on the market – like active power or ancillary service – is involved in organized crime. He tries to eavesdrop on participating agents to get confidential information about processes at the marketplace and uses this confidential information to get financial advantages.
- **Threat scenario 6 (TS6): Gap elimination**
  A unit agent has a technical background and big interests to work with his power plant in a secure environment. So he checks the system using penetration testing toolkits to monitor different parts of the system to eliminate different gaps.
- **Threat scenario 7 (TS7): Denial of Service**
  A grid agent – who evaluates the current and expected grid states in his topology and is responsible for the power quality of his grid – is hijacked by an extrinsic botnet. His availability is threatened and hence, e.g., runtime data will be delayed or even completely interrupted. Thus, an instable grid cannot be detected. As a result, the application of an emergency management plan can lead to power down power plants.

[8] Market agents communicate with coalition agents regarding products on markets.

[9] Grid agents evaluate the current and expected grid states in their topology.

[10] See BSI-Standard 100-3: https://www.bsi.bund.de/DE/Publikationen/BSI_Standard/it_grundschutzstandards.html.

[11] See NIST Special Publication 800-30: http://csrc.nist.gov/publications/nistpubs/800-30/sp800-30.pdf.

- **Threat scenario 8 (TS8): Aimed attack**
  A partner of a unit agent is very similar to the malicious unit agent (e.g., both are combined heat and power plants with comparable output and flexibility). The malicious unit agent tries to get more orders and hence gain more financial capital. For this reason he tries to decrease the trustworthiness of this particular unit agent partner.
- **Threat scenario 9 (TS9): Own upgrading**
  A unit agent tries to manipulate the data in order to increase his own trustworthiness to get more orders and, hence, more profit. This threatens the authenticity of the unit agent and the integrity of the trustworthiness.

	Threatened security goals					Attack motivations						Agent type			Attack origin	
	confidentiality	integrity	availability	authenticity	non-repudiation	social	technical	political	financial	governmental	inadvertent	grid	unit	market	intrinsic	extrinsic
TS1 – Identity fake				X					X				X		X	
TS2 – Unreliable agent			X		X				X		X		X		X	
TS3 – Extrinsic attacker	X	X								X						X
TS4 – Political destruction			X	X				X				X				X
TS5 – Organized crime	X								X					X	X	
TS6 – Gap elimination	X	X	X	X	X		X						X		X	
TS7 – Denial of service			X			X						X				X
TS8 – Aimed attack		X							X				X		X	
TS9 – Own upgrading		X		X					X				X		X	

Table 1: Categorization of identified threat scenarios

The previously described threat scenarios cover the different mentioned categories. Not all threats or attacks can be identified or averted by a trust model but in conjunction with security measures the threats can be minimized. In the project “Smart Nord” the trust model will only be used for the formation of coalitions at the power market. The main threats that have to be covered by the trust model are the financially motivated ones, because marketplaces often have problems with monetary aspects. The security goals which have to be included in the trust model are authenticity and integrity because it has to be ensured that the identity of an agent is authentic and the data is not manipulated – otherwise a trust value is worthless.

Different types of malicious agents from the threat scenarios TS1, TS2, TS5, TS8 and TS9 will be implemented to evaluate the trust model. The other threat scenarios will be only considered in the development process of the system to avoid security gaps, but not implemented.

4.3 Evaluation questions and criteria

As mentioned in Section 4.2, some types of malicious agents with different goals have to be implemented in the “Smart Nord” project to enable the evaluation of the trust model. The following exemplary questions have to be answered in this evaluation.

- After which number of coalition formation iterations does the tampering or the misbehaving of an agent have to be revealed?
- How presumable is the tampering of the trust value in a special situation (e.g. the higher the capital the higher the misuse probability)?
- How many recommendations have to be considered to get a useful opinion of a foreign agent?

There are a lot more evaluation questions which have to be developed in greater depth in the future work.

5. Conclusion and further work

In this paper first results of modeling a trust model were described. First, a definition of trustworthiness was given. After that an exemplary scenario of a trustworthy coalition formation in the project context is shown. The main results of this paper are nine threat scenarios for multi-agent systems that are related to different security goals and attack motivations. These threat scenarios will be used for the evaluation of the trust model in the context of the mentioned project “Smart Nord”. For this evaluation, some questions are introduced as scientific goals.

For future work, the implied trust model has to be defined explicitly and the security aspects in that model have to be specified to secure the processes. Additionally, the exemplary scenario has to be enhanced to assemble it with the threat scenarios and to use them for the trust model evaluation.

Acknowledgements

The Lower Saxony research network 'Smart Nord' acknowledges the support of the Lower Saxony Ministry of Science and Culture through the “Niedersächsisches Vorab” grant programme (grant ZN 2764).

Bibliography

Blank, M.; Lehnhoff, S.: Assessing Reliability of Distributed Units with Respect to the Provision of Ancillary Services, Proceedings of 11th IEEE International Conference on Industrial Informatics, INDIN 2013, 2013.

Geschonneck, A. (2008): Computer Forensik – Computerstraftaten erkennen, ermitteln, aufklären, ISBN 978-3-89864-534-8, dpunkt.verlag.

IEC (2007): IEC 62351-7 TS Ed. 1: Power systems management and associated information exchange - Data and communication security - Part 7: Network and system management (NSM) data object models".

Jøsang, A., Ismail, R., Boyd, C. (2007): A Survey of Trust and Reputation Systems for Online Service Provision, Article published in Decision Support Systems, 43(2) 2007, p.618-644.

Kiefhaber, R., Satzger, B., Schmitt, J., Roth, M. and Ungerer, T. (2010): Trust Measurement Methods in Organic Computing Systems by Direct Observation, Proceedings of 2010 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, IEEE computer society.

Rosinger, C., Uslar, M., Hockmann, F. (2012): Reputationssysteme für selbstorganisierte Multi-Agenten-Systeme in Energiemanagementsystemen, VDE-Kongress 2012, ISBN 978-3-8007-3446-7.

Sonnenschein, M., Appelrath, H.-J., Hofmann, L., Kurrat, M., Lehnhoff, S., Mayer, C., Mertens, A., Uslar, M., Nieße, A., and Tröschel, M. (2012): Dezentrale und selbstorganisierte Koordination in Smart Grids, VDE-Kongress 2012, ISBN 978-3-8007-3446-7.

Steghöfer, J.-P., Kiefhaber, R., Leichtenstern, K., Bernard, Y., Klejnowski, L., Reif, W., Ungerer, T., André, E., Hähner, J. and Müller-Schloer, C. (2010): Trustworthy Organic Computing Systems: Challenges and Perspectives, Proceedings of the 7th International Conference on Autonomic and Trusted Computing, Springer.