Practical Use Case Evaluation of a Generic ICT Meta-Risk Model Implemented with Graph Database Technology

Stefan Schiebeck, Martin Latzenhofer,
Brigitte Palensky, Stefan Schauer
Digital Safety & Security
Austrian Institute of Technology
Vienna, Austria
e-mail: {stefan.schiebeck.fl | martin.latzenhofer |
brigitte.palensky | stefan.schauer}@ait.ac.at

Gerald Quirchmayr
Research Group Multimedia Information Systems
Faculty of Computer Science
University of Vienna
Vienna, Austria
e-mail: gerald.quirchmayr@univie.ac.at

Thomas Benesch
Research & Development
s-benesch
Vienna Austria
e-mail: thom@s-benesch.com

Johannes Göllner, Christian Meurers, Ingo Mayr
Department of Central Documentation & Information
National Defence Academy of the Austrian Federal
Ministry of Defence and Sports, Vienna, Austria
e-mail: {johannes.goellner | christian.meurers |
ingo.mayr}@bmlvs.gv.at

***Abstract*— Advanced Persistent Threats impose an increasing threat on today's information and communication technology infrastructure. These highly-sophisticated attacks overcome the typical perimeter protection mechanisms of an organization and generate a large amount of damage. In this article, we introduce a generic ICT meta-risk model implemented using graph databases. Due to its generic nature, the meta-risk model can be applied on both the complex case of an APT attack as well as on a conventional physical attack on an information security management system. Further, we will provide details for the implementation of the meta-risk model using graph databases. The major benefits of this graph database approach, i.e., the simple representation of the interconnected risk model as a graph and the availability of efficient traversals over complex sections of the graph, are illustrated giving several examples.**

***Keywords*— risk management; APT; ICT security; physical security; graph databases; interconnected risk model.**

I. INTRODUCTION

Based on a practical use case of a real-life Advanced Persistent Threat (APT) lifecycle, we showed in a recent article [1] how this type of attack can be tackled by a generic information and communication technology (ICT) meta-risk model using graph databases. In the present article, we will extend our preliminary work and show how the meta-risk model can be applied in a different context, i.e., a physical attack scenario.

Although internal attacks can be seen as today's biggest threat on information security [2], in practice, information security officers still put great emphasis on perimeter control. The internal area of a company's ICT network, e.g., the demilitarized zone (DMZ) or the intranet, is secured based on standard technical guidelines demanding, e.g., the logical separation of a network into subnetworks according to specific security requirements [3]. Nevertheless, the effort invested in monitoring the internal network is moderate. Intrusion detection and prevention systems are cost and time consuming and require a large amount of administration. Recent attack strategies like APTs take advantage of this lack of internal control.

The term APT summarizes a family of highly sophisticated attacks on an ICT network or infrastructure. Usually, an APT runs over an extended period of time with the objective to steal data and maintain presence indefinitely without being detected. A continuous access allows collecting new data as it emerges, extending the achieved foothold over time, and using the site as a jumping point for the attack on other facilities. The adversary – usually a group of people – has a large amount of resources at hand and applies the whole range of digital, physical and social attack vectors to gain access to a system. The attack is specifically designed for a particular victim, i.e., a company or an organization, such that common security measures can be circumvented effectively. Thus, the adversary stays undetected over a long period of time. One particular technique recurrently used in APTs is social engineering, which exploits the human factor as a major vulnerability of an ICT system. Potential countermeasures, like increasing the staff's awareness concerning ICT security threats via training courses, are not very common. According to a Ponemon study [4], about 52% of the interviewed organizations do not offer respective training courses for their employees.

In the course of the last decade, APTs became one of the most significant kinds of threats on information security, causing a great number of security incidents all over the world [5]. Besides the most prominent APT attack, the application of the malware Stuxnet in an Iranian nuclear

power plant, a number of other APT attacks have become known, e.g., Operation Aurora, Shady Rat, Red October or MiniDuke [6][7][8]. As it is shown in the Mandiant Report [5], some adversaries even have a close connection to governmental organizations. The former director of the US cyber command, General Keith Alexander, referred to the currently occurring industrial espionage and theft of intellectual property as ”the greatest transfer of wealth in history” [9]. In Europe, the disclosures of Edward Snowden [10] have drawn great attention to this issue. Based on current numbers from cyber-crime reports, which show the growing amount of damage [11][12], it is distressing how poorly evolved today’s countermeasures seem to be.

This article focuses on the implementation of a generic ICT meta-risk model that can deal both with the described issues and can be applied on conventional ICT security use cases as well – e.g., a physical attack on a building with the aim to gain access to some information. The implementation of the meta-risk model is based on graph databases and social network analysis concepts to provide a perspective that can focus on a specific aspect (node) and its influences (relationships). From a technological perspective, the advantages of the chosen approach are demonstrated, in particular concerning aggregation of exposures, risks, etc.. Therefore, different types of assets, e.g., organizational aspects like processes and personnel, ICT components like IT systems and logical networks, and physical infrastructure objects, serve as examples of assets that are attacked in fictitious, but realistic ways.

In detail, after a short overview of related work on graph-based models in Section II, Section III sketches the different steps of an APT attack for a fictitious scenario to illustrate the basic principles of this family of threats. Section IV introduces the theoretical background and the development of the generic ICT meta-risk model depicted as a graph model. The subsequent Section V shortly discusses the pros and cons of an implementation via graph databases vs. relational databases. Sections VI and VII show the modeling of the two use cases introduced in this article, the APT and the physical attack scenario. Section VIII provides a detailed description of how the generic risk model is implemented using a graph database. Finally, Section IX summarizes the results.

II. Related Work

Whereas the internationally widely spread ISO 31000 standard [13] provides generic guidelines for the design, implementation and maintenance of risk management processes throughout an organization, the ISO/IEC 27005 standard [14] specifically focuses on information security risk management. In [15], this standard is taken as a basis and extended by the introduction of iteratively calculated management measures, indicators and expert knowledge, as well as the possibility to integrate sensors for automation purposes. The resulting continuous information security risk management model (cf. Figure 1) is also demonstrated and verified by a prototype and provides a framework for extendable sensors. This framework can be used to continuously gather security relevant attributes having a high impact on the overall model, derive security metrics and indicators based on ISO/IEC 27004 [16] and enable adaptable knowledge management approaches to infer risk factors of a risk assessment model. In the KIRAS project MetaRisk [17], the approach of [15] is connected with meta-models for organization planning and control to derive a comprehensive enterprise risk management system referred to as meta-risk model. Additionally, a graph-based implementation has been introduced, which allows the visualization and semi-heuristic handling of complex relationships in a schema-free form.

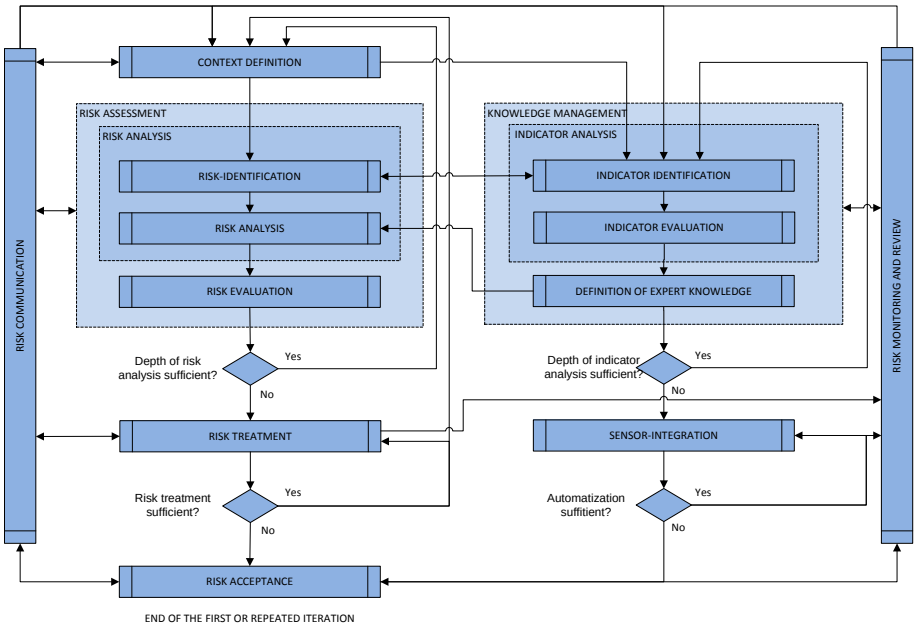

Figure 1. Extended ISO/UEC 27005 risk management process.

In general, graph-based models are used to capture relations among system entities at various abstraction levels. In [18], Chartis Research advises the introduction of graph analytics (based on graph databases) to the risk management activities of financial institutions so that they can discover so far unknown risks by revealing interconnected risk patterns. In [19], graph-based representations are applied in the area of risk management for critical infrastructures (CI). Bayesian Networks are used to learn (or simply estimate) CI service risks and their interdependencies. Additionally, a risk prediction is introduced and a case study to validate the model is carried out. However, some of the model's features, like risk prediction and the handling of cyclic dependencies, could not be verified because they simply did not occur during the run-time of the case study. In contrast to the work presented in this article, the main goal of the approach in [19] is to identify an abstract set of variables and their dependencies based on system measurements and it is for this purpose that graph-based representations are introduced. The direct use of graph databases is not foreseen or discussed in [19].

The continuous risk management process depicted in Figure 1 essentially consists of four main steps, which are performed iteratively (cf. in particular with the additional parts to the standard version on the right side of Figure 1 referred to as Knowledge Management). These steps are the business value analysis, the scenario analysis, the threat analysis and the relation analysis (further described in Section IV. All risk-relevant measures and events available within an organization can be integrated in the continuous risk assessment process by repetitively identifying these measures (categorized and compressed according to ISO/IEC 27004 [16]) and evaluating them with respect to indicators. The input of formally defined expert knowledge ensures a continuous update of the used risk factors and their corresponding indicators. The development and application of sensors for an ongoing collection of relevant data leads to a higher degree of automation.

In this article, we introduce the explicit usage of graph databases and combine it with the aforementioned, already existing risk scheme retrieved mainly from the IT-Grundschutz catalog described in [15]. This approach has initially been shown in [1] and is extended here with a conventional use case implementation for physical information security. However, many important extensions have been made to this scheme to derive a much more generic model. The presented approach enables a measurable iterative increase of the depth of the risk analysis on all the analysis levels as well as an improved risk treatment. It enables the setup of an appropriate balance between required effort and obtained risk coverage. Furthermore, using the approach, cascading risks can be represented in a straight-forward way that allows us to run easily through a typical APT attack scenario. The underlying model and functional assessment concept presented in this article, excluding the usage of a graph database for data manipulation, have been demonstrated in [20], although with the use of a relational database.

Before going into more detail on the description of the meta-risk model in Section IV and its implementation in Section V, we will present the general APT use case scenario we will rely our further discussion on.

III. APT SCENARIO

In [5], the US security company Mandiant describes the typical lifecycle of an APT attack based on an analysis of how a Chinese cyber espionage group infiltrated several companies in the US and worldwide. In the following, the different steps in this APT lifecycle are briefly sketched to give an overview on the basic operations of an APT attack (cf. Figure 2). To provide a better illustration of the scenario, a fictional research facility, *Biomedical Research*, is used. It consists of four research laboratories with increasing degrees of security requirements (*Biosafety Level 1-4*) located in physically separated buildings. Additionally, the research facility runs two data centers, one located in the research building itself, and the other, which works as a backup, located at a distant administrative building. The information most valuable for an attacker is assumed to be hosted in Research Laboratory FL4, which is the one with the highest security level, or in one of the data centers. Based on this setting, a generalized APT attack can be outlined in eight steps.

As a first step, *Initial Recon*, the adversary tries to gain access to the organization's ICT infrastructure. Since the terminals in Research Laboratory FL1 are the only ones having full connection to the internet, a user in FL1 would be a primary target for a spear phishing attack (cf. (1) in Figure 1) in order to place a remote backdoor on either of these terminals. A potential user to be attacked can be identified for example using social engineering. In the second step, *Initial Compromise*, a user in FL1 receives a spear phishing mail and opens the infected attached file (e.g., a ZIP-file). During the execution of the ZIP-file, a basic backdoor (beachhead backdoor, cf. (2) in Figure 2) is installed on the terminal W1. Through this backdoor, a connection to the adversary's command and control server is established. In a third step, *Establish Foothold*, this initial connection is used to install a standard backdoor on the compromised terminal, giving the adversary an increased set of possibilities. Hence, the adversary is able to gain foothold at the application server S1 in FL1 (cf. (4) in Figure 2).

The following four steps (steps 4 to 7) are usually performed more than once, until the adversary acquires the desired information. In step 4, *Escalate Privileges*, the adversary gathers information on valid combinations of user names and passwords inside the internal networks. The attacker also gains additional information about the internal network structure (step 5 – *Internal Recon* – cf. (5) in Figure 2), potentially including internal authentication information. In the following step 6, *Move Laterally*, the adversary infiltrates the local data center as well as the backup data center to locate the valuable information. This is achieved using a vulnerability scan on the file servers S7.1 and S7.2 and an appropriate exploit allowing the compromise of both identically configured systems (cf. (6) in Figure 2). As a final step of this recurrent loop, *Maintain Presence*, all tracks are

covered up and the adversary silently stays in the victim's system with an extended foothold (cf. (7) in Figure 2).

The final step, *Complete Mission*, starts when all the target information is collected. Covert channels are established (e.g., using cryptography/steganography) to extract the sensitive information from the file servers (cf. (8) in Figure 2). Afterwards, all traces of the attack are erased.

IV. Setup of the ICT Meta-Risk Model

The risk analysis process of the continuous information security risk management model presented in [15] (cf. Section II above) incorporating the knowledge management parts comprises of the following layers or steps:

- *business value analysis* for the systemic representation of all the assets that need to be protected,
- *scenario analysis* (optional) for the representation of high-level dependencies between assets,
- *threat analysis* (optional) for the specific modeling of low-level threat cascades,
- *relation analysis* (automatic) which delivers a combined risk overview over all the modeled scenarios for each asset.

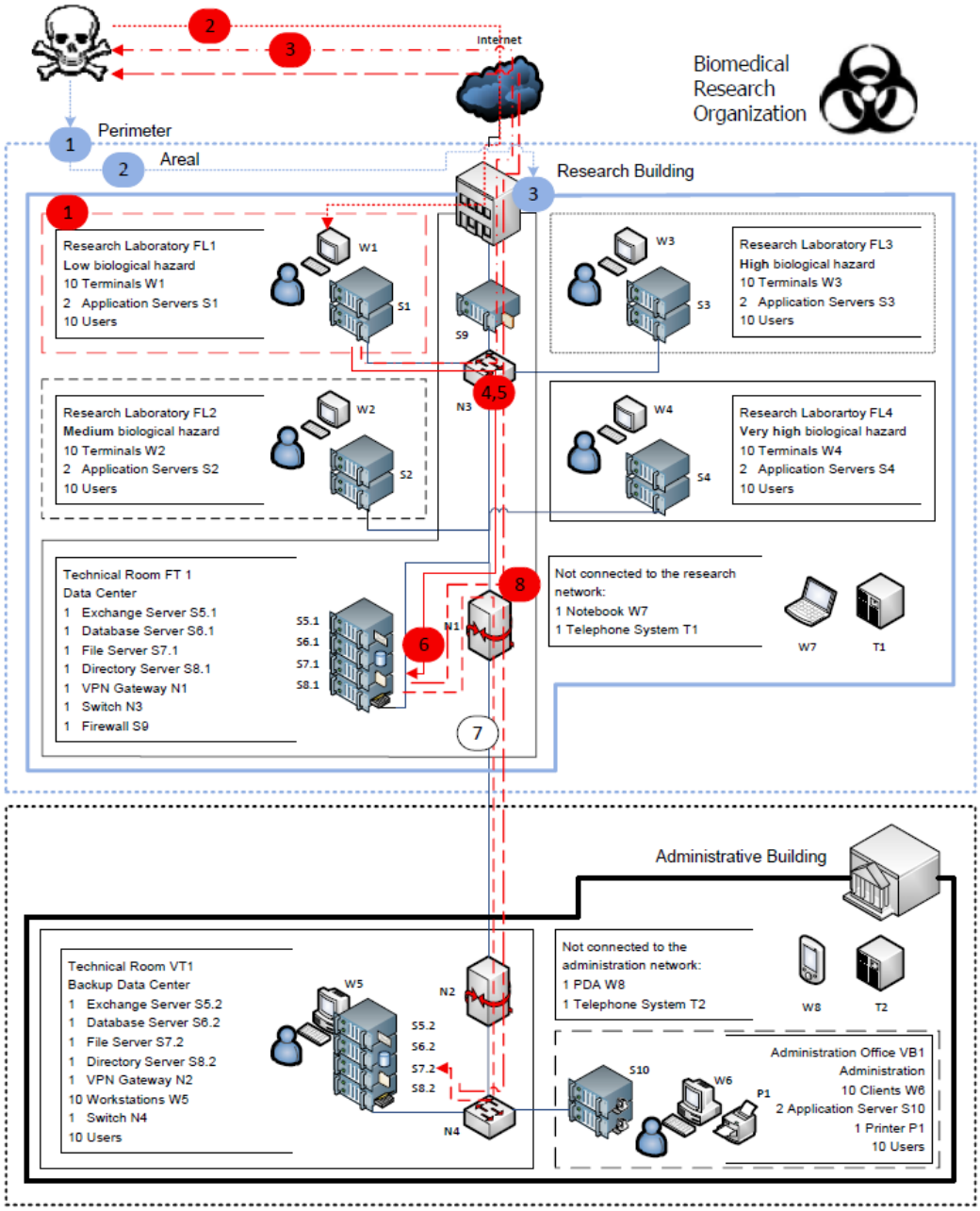

Figure 2. Sample scenarios: APT attack (red), physical security attack (blue).

The Business value, scenario and threat analysis can be modelled with iterative increase of modelling depth and detail, while the relation analysis automatically incorporates all asset instances modeled. So far, the implementation of the prototype of the proposed general model and thus also of the risk-analysis is largely based on the standards, catalogues and cross references from the BSI's IT-Grundschutz [15]. The advantage of using the IT-Grundschutz approach is that it delivers an extensive list of IT-related threats, which are already connected with assets, together with safeguards against these threats and roles that are responsible for planning and implementation of these safeguards. Although using it as a basis, our model extends the BSI approach in several aspects, e.g., concerning the view on the protection criteria (i.e., confidentiality, integrity and availability), or the introduction of risk management aspects. In the following, we will describe the four steps of the risk analysis process of our model (cf. also Figure 3).

A. *Business Value Analysis*

In the first step, all assets of an organization requiring protection have to be identified. This can be done, for example, using the IT-Grundschutz catalog. In this case, the business assets are represented by one or more modules from the IT-Grundschutz. In this context, standard assets are, for example, applications, IT-systems, networks, rooms, and buildings. Additionally, in more complex models also legal entities, organizational divisions, or processes can be taken into account by representing them in the form of modules. In any case, each module has several protection criteria (e.g., confidentiality, integrity, and availability) and is associated with various threats, which, in turn, are related to protection criteria on the one hand, and appropriate security measures to mitigate them on the other hand (cf. Figure 4).

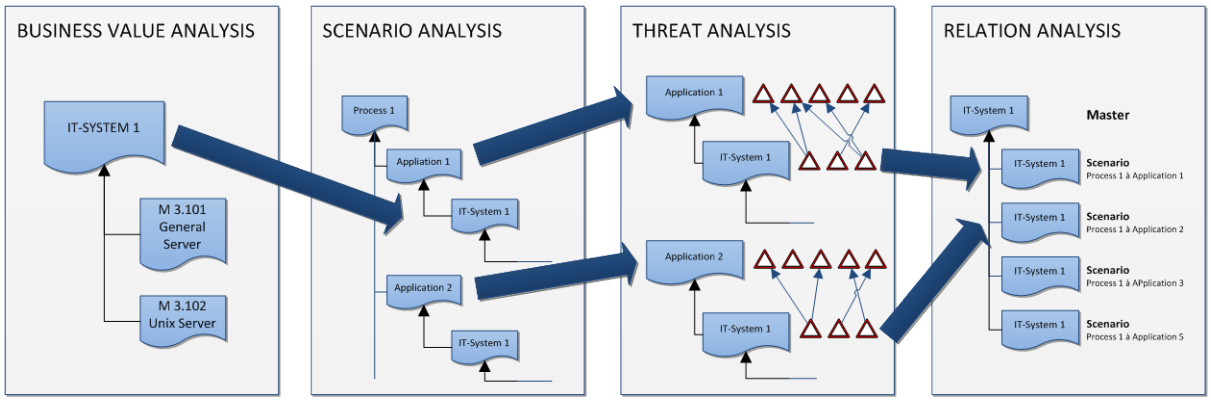

Figure 3. Analysis layers of the overall model.

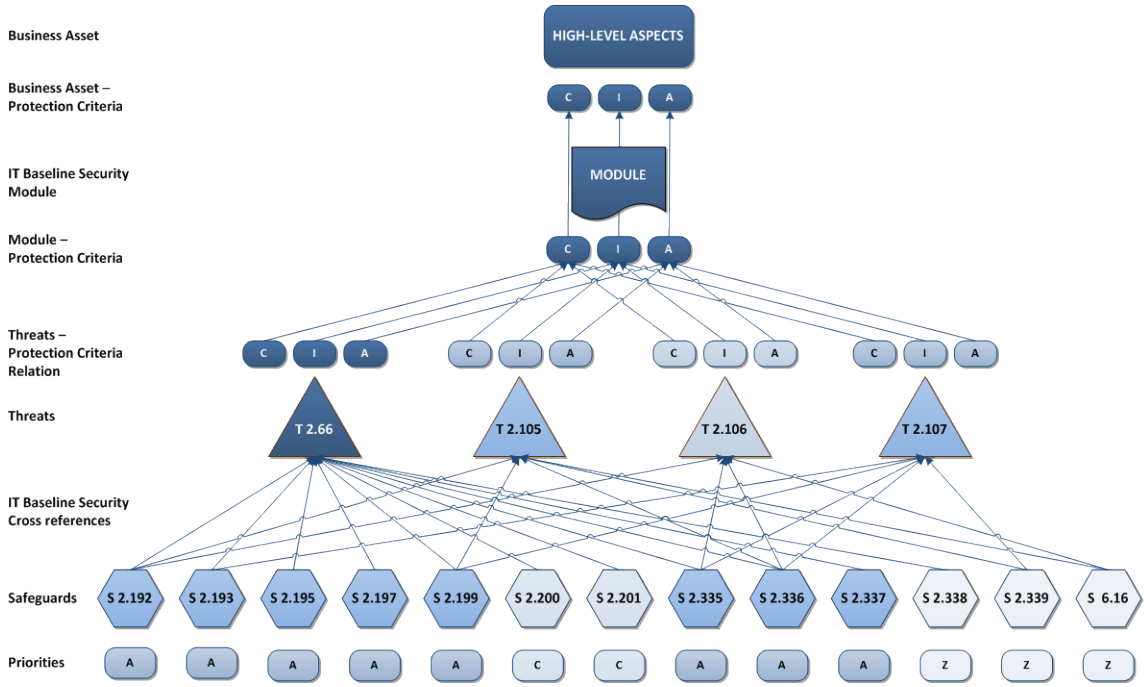

Figure 4. Business Asset Analysis.

The information contained in the interrelations between modules, threats, and security measures is used for the calculation of a value for the exposures of an asset. Thereby, the threat exposure is a function of the likelihood of an attack and the vulnerability of the threatened asset (which largely depends on the maturity levels of the related security measures). Higher level exposures (e.g., module exposure, asset exposure) are aggregated using generic estimation functions (e.g., maximum, sum, energetic sum, etc.). At the asset level, each protection criteria (e.g., confidentiality, integrity, availability, etc.) has an associated requirement level which corresponds with the maximum tolerable impact. By multiplication of protection criterion impact of an asset with its distinct exposure, , an asset risk value is obtained for this protection criterion (e.g., availability risk). An overall asset risk value can, for example, be estimated based on the sum of its protection criteria risks.

B. *Scenario Analysis*

In the scenario analysis [21], the identified business assets are connected with each other. This is done in a structural way, starting from high-level assets (e.g., legal entities or business processes) and going down to more and more specific assets they depend on (e.g., applications, IT systems and networks, buildings and personnel) (cf. Figure 5). Based on the determined dependencies between assets the necessary risk inheritance functions between assets can be set up. Moreover, with the obtained structural knowledge, a business impact analysis [22] can be carried out to identify the protection criteria requirements of each of the assets in various scenarios. In the course of the business impact analysis, a choice of inheritance functions for the protection criteria associated with assets also takes place.

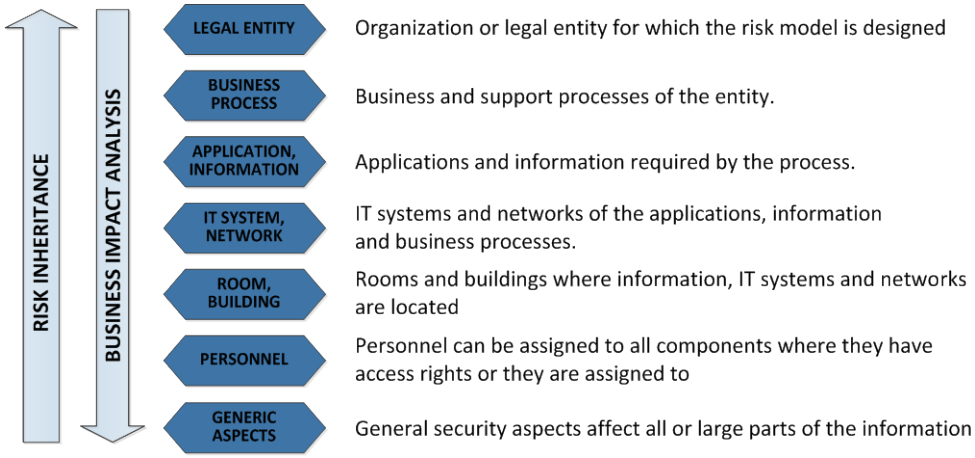

Figure 5. Generalized structure of a scenario analysis.

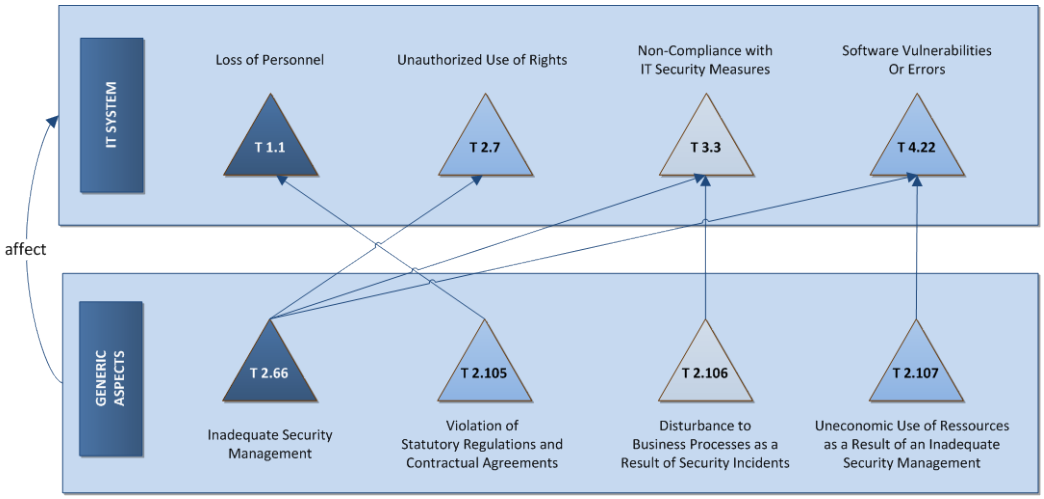

Figure 6. Representation of threat cascades.

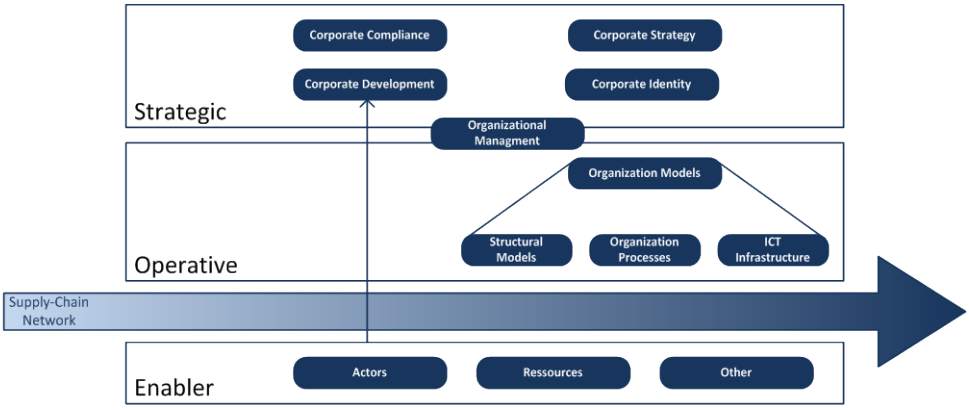

Figure 7. Meta-model of an organization [23].

C. Threat Analysis

Building on the aforementioned structural relations describing the risk inheritance between different business assets, a representation of how different threats affect each other can be obtained. Figure 6 schematically presents several such threat cascades between two assets. In detail, the first asset is an ICT system ("Generic Server") and the second one is an organizational entity ("Security Management") [24]. From this representation it is easy to see that an "Inadequate Security Management" (T2.66) can lead to several other threats, i.e., "Unauthorized Use of Rights" (T2.7), "Non-Compliance with IT Security Measures" (T3.3), and "Software Vulnerabilities or Errors" (T4.22). By iteratively extending the model of low-level threat cascades between the assets connected during scenario analysis, the estimation model gains knowledge about adapted threat exposures based on required threat predecessors.

D. Relation Analysis

As a final step, to get an estimation of the threat exposure per asset additionally to the scenario-based threat picture, all the risk carrying scenarios that affect a specific asset are aggregated. Further, the protection criteria values coming from the various scenarios are combined per asset. The negative effects of unwanted events on the protection criteria of an asset are a measure of their impact on that asset. By combining the threat exposure with the impact, a risk value can be calculated.

E. ICT Meta-Risk Model

In the course of the MetaRisk project, the approach described above has been further extended to develop a comprehensive ICT meta-risk model. Derived in a combined bottom-up/top-down way, this generic ICT meta-risk model subsumes all the typical components of common risk management models, tools, processes, and control logic. Its central component is a meta-model of an organization, which aims at describing an organization in a holistic way (cf. Figure 7 for a schematic representation and [23] for a more detailed description of the model).

This meta-model of an organization builds upon three layers: a strategic layer, an operative layer, and a layer that subsumes the enablers of the organization. It also includes the organization's supply chain network. On the strategic level, the overall strategy of the organization is described, together with the corporate compliance, the corporate identity and the development of the organization. These four topics influence the risk management on the highest level, defining the long-term goals of the organization.

On the operative layer, more detailed models can be found. These include the organization's structural models, i.e., the general setup of the organization, including buildings, machines and other tangible objects, as well as process models describing the activities and day-to-day business of the organization. A special focus is laid on the ICT infrastructure of the organization, since it represents a core feature of every organization and is crucial to achieve the organization's goals.

The third layer describes the enablers, i.e., all the actors and resources required to perform the organization's daily business. In this context, actors are usually divided into several groups and types of actors, often specific to the underlying organization, together with their respective roles according to knowledge management. The enablers have to be seen as key factors in the overall risk management process since they can represent threats, targets, and safeguards (i.e., mitigation actions).

Starting from the described generic model of an organization, several standard processes and frameworks for risk assessment and risk management are combined in a bottom-up approach to derive the ICT meta-risk model. Therein, the plan-do-check-act (PDCA) cycle defined in the ISO 31000 [13] represents the reference for the basic process and categorization model. A detailed analysis of several further standards and frameworks has shown that the PDCA cycle works as a robust basis for their integration. Generic modelling requirements outlined in ISO 31000, ISO 27000, ISO 28000 and, e.g., OCTAVE have been used to perform completeness checks on the ICT meta-risk model. Furthermore, several frameworks, primarily the IT-Grundschutz catalogues, COBIT 5 as well as the respective control mappings and goal cascade information, have been integrated as modelling catalogs.

The resulting ICT meta-risk model (cf. Figure 8) can be represented by a graph. It integrates all information required for the modeling and computation of risk objects within an organization. The intended purpose of the generic graph-based ICT meta-risk model is to provide an easy-to-extend and schema-less representation with the ability to interrelate different types of nodes and to aggregate information across affected relationships.

V. PROTOTYPE IMPLEMENTATIONS

In the following, we will describe the implementation details of the ICT meta-risk model in graph databases. The underlying approach covers semi-quantitative analysis steps usually used within risk models applying ISO 27005 [14]. We focus on the interconnections in the graph-based meta-model (cf. Figure 8), their representation in the graph database, and the relation to the implementation of the APT attack scenario and the physical attack scenario therein.

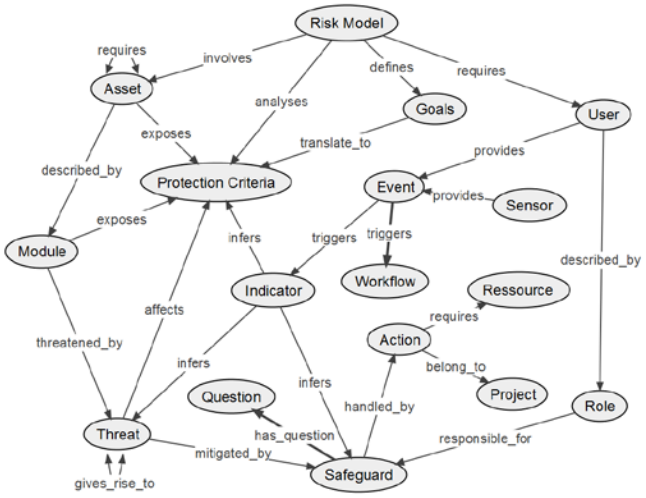

Figure 8. Graph-based meta-risk-model.

A. Graph Databases

In this work, as architectural background for the implementation of the model, the graph database Neo4j [25] is used instead of a relational database. Graph databases provide the advantage of being able to perform near-real-time traversals and aggregations, efficient topology analyses, and the optimal finding of node neighbors [26]. The retrieval time of graph databases is usually significantly less than that of relational databases [27][28]. Moreover, the graph-based implementation ensures more flexibility for defining relationships between datasets. Whereas relational databases are difficult to extend, in graph databases only a few edges and nodes have to be added to extend the graph. Thus, the adaption and extension capabilities of the generic ICT meta-risk model are supported by the schema-less definition of data in graph databases. For instance, additional information on customer and competitor intelligence, responsibilities, or other quantifiable business data can be easily integrated in the database schema. When using Neo4j, the integrated declarative query language CYPHER [25] supports most of the work. Thus, analysis models with extended and adapted functionality are greatly simplified and the graph-based approach is more efficient than business code migration on the software backend.

In situations, where the data set is quite homogenous and rarely changed, other architectural designs, like relational databases or in-memory databases, may be more appropriate. They also offer more support as well as advantages in the field of maturity. Regarding security, MySQL has an extensive security support based on access control lists. In contrast, graph databases like Neo4j expect a trusted environment.

B. Graph-based ICT Meta-Risk Model

The starting point for the graph-based ICT meta-risk model is the node *Risk Model*. This risk model contains narrative information on the scope of the risk analysis, the goals as well as its requirements. Each risk model has several goals relevant for the analysis attached to it (*Risk Model defines Goals*). For the categorization of these goals we will use the taxonomy coming from the IT-Grundschutz. In detail, the following categories are distinguished:

- *Modules*: applications, IT systems, networks, infrastructure, high-level aspects, etc.
- *Threats*: elementary risks, force majeure, organizational deficiencies, human failure, technical failure, intentional actions, etc.
- *Safeguards*: infrastructure, hard- and software, emergency planning, organization personnel, communication, etc.

The integration of several frameworks into the ICT meta-risk model allows us to use the taxonomy of COBIT 5.0 and to deduce COBIT-specific goals, e.g., stakeholder needs, enterprise goals, IT-related goals. These goals correlate to the exposure of the components in the risk model and therefore affect the relevant protection criteria, e.g., confidentiality, integrity, availability (*Goals translate_to Protection Criteria*). Furthermore, the ICT meta-risk model also allows evaluating these protection criteria separately (*Risk Model analyses Protection Criteria*).

Users can be associated to the risk model (*Risk Model requires User*) in specific roles (*User described_by Role*) regarding the planning, implementation, and audit of required safeguards (*Role responsible_for Safeguard*). Users as well as automatable sensors using pre-aggregated data from external support systems (e.g., security information management solutions or security incident and event management (SIEM) systems) can provide measurements and events to the framework (*User/Sensor provides Event*). Events can be used to trigger workflows (*Event triggers Workflow*), e.g., when a new IT system is detected. The framework also provides a possible inference option between objective measurements and related subjective risk factors using fuzzy indicators (E*vent triggers Indicator*) and an expert knowledge system. There might be the following interferences:

- Protection criteria (*Indicator infers Protection Criteria*) meaning that the indicators refer to the estimated damage

- Threats (*Indicator infers Threat*) meaning that the indicators refer to probabilities of occurrence
- Safeguards (*Indicator infers Safeguard*) meaning that the indicators refer to the exploitation potential of vulnerabilities.

In context of some events it can be useful to trigger workflows (*Event triggers Workflow*), e.g., integrating new vulnerabilities of IT systems into the risk model, which were discovered during scans. In order to support basic risk management functionalities, safeguards can be summarized as organizational actions (*Safeguard handled_by Action*), which can be combined with resources, e.g., personnel, finance, etc. (*Action requires Resource*) or projects (*Action belong_to Project*). By integrating priorities, return on security investment models can be feed with the results from cost and availability information analyses.

Pre-existing information including goals, boundaries, and requirements are narratively documented within the nodes of the risk model. Risk identification is carried out by the definition of organizational assets (*Risk emergency planning,Model involves Asset*) that are depicted by modules (*Asset described_by Module*), threats (*Module threatened_by Threat*), safeguards (*Threat mitigated_by Safeguard*), and roles (already introduced: *Role responsible_for Safeguard*). Goals can be defined based on the usual protection criteria (confidentiality, integrity, availability), as well as on requirements derived from other taxonomies. IT-Grundschutz [15] defines a respective risk catalog, providing a categorization by module type (applications, IT systems, networks, infrastructure, common aspects), threat type (basic, force majeure, organizational shortcomings, human error, technical failure, deliberate acts), and safeguard type (infrastructure, organization, personnel, hardware and software, communications, contingency planning). Moreover, additional goals and requirements (e.g., stakeholder needs, enterprise goals, IT-related goals, etc.) coming from different frameworks like COBIT [29] can be integrated using cross-references with IT-Grundschutz. The defined goals correlate with the respective exposure of the components within the risk model, which translate to several risk dimensions *Risk Model analyses Protection Criteria*).

Risk estimation is based on the determination of safeguard maturities (supported by additional control questions, *Safeguard has_question Question*), threat likelihoods, and impacts on protection criteria. As a result of estimation, exposures are calculated for assets, modules, and threats separately (*Asset/Module exposes Protection Criteria; Threat affects Protection Criteria*) (cf. Section 0).

Assets can optionally be related to each other during scenario analysis in order to depict their dependencies (*Asset requires Asset*). This supports business impact analysis and the option to perform risk propagation between scenario assets. Another optional step is to perform a detailed threat analysis by modeling threat cascades [30] (*Threat gives_rise_to Threat*) based on the relationships of the pre-structured scenario model (*Asset requires Asset*).

VI. Modeling the APT Scenario

In the following, the graph-based model of the APT use case scenario described in Section III is discussed in detail (cf. Figure 9). Assets (blue ovals) are modeled by *_requires_* dependencies, which can be identified by a scenario analysis. The resulting structure defines the top-down inheritance between sub-systems and, at the same time, serves as default path for potential bottom-up threat cascades (*_gives rise to_*). Assets are connected to IT-Grundschutz modules (yellow hexagons) [15], where the referring relation is *described_by*. Threats (red trapezia) are linked to assets by *threatened_by* relations and associated with security measures (green rectangles) by *mitigated_by* relations. For the purpose of a detailed analysis, available threats can be combined to threat cascades via *gives_rise_to* relations. The business impact analysis model (*described_by*) and the IT-Grundschutz taxonomy itself indicate how these cascading paths might look like. This approach of modeling cascades might not address all of the potentially existing correlations, but it provides an easy way of dealing with chained probabilities.

When looking in detail at the APT attack scenario as described in Section III, we see that initially a user opens a spear phishing mail at the Terminal W1 (*Module M 3.201 General client*). This is an exploitation of the organizational threat *T 3.3 Non-compliance with IT security measures*, which is connected with the following security measures:

- *S 2.23 Issue of PC Use Guidelines*
- *S 4.3 Use of virus Protection Programs*
- *S 4.41 Use of appropriate security products for IT systems*

Afterwards, at the corresponding terminal server S1 (*Module M 3.305 Terminal servers*) a standard backdoor is installed. This is possible because of the threat *T 2.36 Inappropriate restriction of user environment*, which could have been addressed by the following security measures:

- *S 2.464 Drawing up a security policy for the use of terminal servers*
- *S 4.365 Use of a terminal server as graphical firewall*
- *S 4.367 Secure use of client applications for terminal servers*

Having gained access to the Terminal server S1, a software vulnerability scan is performed, helping the attacker to exploit the threat *T 4.22 Software vulnerabilities or errors* at the File server S 7.1 and, later on, at the file server S 7.2 (*Module 3.109 Windows Server 2008*). In the analyzed use case, the following security measures were not properly implemented:

- *S 2.32 Establishment of a restricted user environment*
- *S 2.491 Use of roles and security templates under Windows Server 2008*
- *S 4.417 Patch Management with WSUS under Windows Server 2008 and higher*
- *S 4.419 Application control in Windows 7 and higher by means of AppLocker*

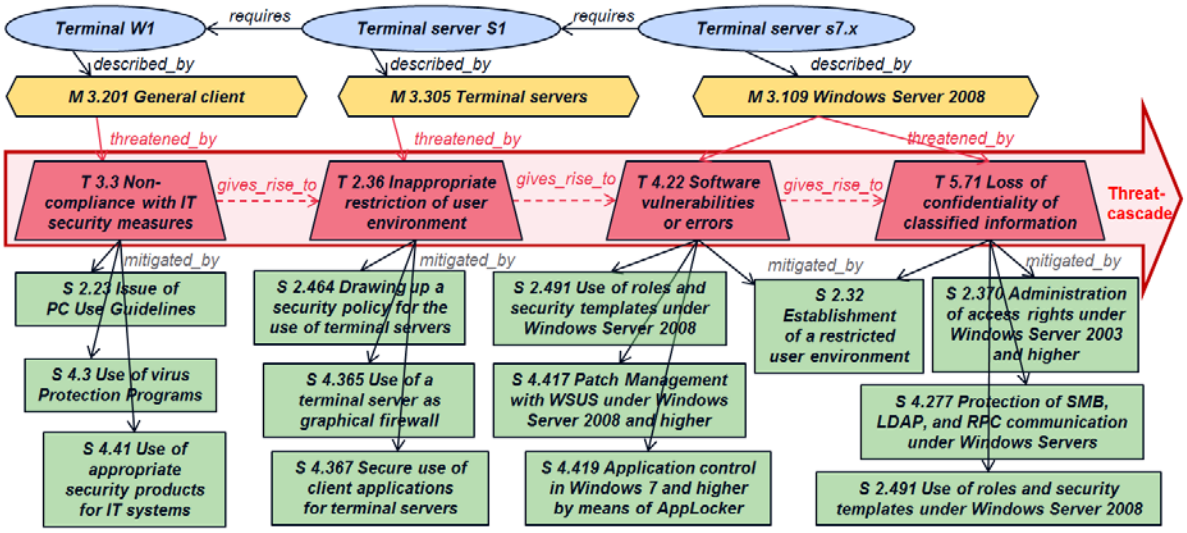

Figure 9. Graph-based illustration of the APT scenario.

At the same module, the follow-up threat *T 5.71 Loss of confidentiality of classified information* can be triggered, which is addressed by the following security measures:

- *S 2.32 Establishment of a restricted user environment*
- *S 2.370 Administration of access rights under Windows Server 2003 and higher*
- *S 2.491 Use of roles and security templates under Windows Server 2008*
- *S 4.277 Protection of SMB, LDAP, and RPC communication under Windows Servers*

In order to perform quantitative analyses, the risk inheritance between different components can be modeled by appropriate functions, e.g., maximum, sum, product, or minimum. More complex normalized, weighted, or bounded variants are also applicable. Possible candidates for the latter are weighted weakest link or prioritized sibling [20][31].

VII. Modeling the Physical Scenario

In this section, it is shown that in an analog way to the ATP attack example, the ICT meta-risk model can also be used to describe and evaluate a physical attack scenario. Therefore, the use case discussed below models a typical physical environment of an ICT infrastructure. This scenario is also depicted in Figure 2. A layered architecture is assumed, i.e., the relevant ICT infrastructure is located within a building, the building is located on the grounds of a company, and the company is protected by a specific perimeter.

A graph-based illustration of this scenario, similar to the one of the APT scenario in Figure 9, is given in Figure 10. Therein, the relevant assets are represented by blue ovals and modeled by *_requires_* dependencies. The associated modules (yellow hexagons) do not correspond to modules from the IT-Grundschutz anymore, but can still be treated as such by the ICT meta-risk model. Thus, each module is linked to threats (red trapezia) by the *_threatened_by_* relations, and threats themselves are linked to mitigation actions (green rectangles) using the *_mitigated_by_* relation.

In the physical security use case, just as in the APT use case, the ICT meta-risk model allows the representation of threat cascades using the *_gives_rise_to_* relations. Thus, the cascading effects of a threat as well as different attack variants affecting various assets (physical objects) can be modeled. Accordingly, not only the analysis of unrelated individual risks of single objects can be achieved but also that of threat cascades and the risks of whole attack chains.

As shown in Figure 10, an attacker who wants to enter a building has to overcome the perimeter protection first. This can be mitigated, for example, by the following safeguards:

- *Protection against climbing over* (e.g., using a barbed wire on top of a fence)
- *Patrols* (i.e., security guards walking around the area to spot trespassing

If no physical barrier is present, an attacker could simply *cross the perimeter* line. This could be mitigated for example by the security measures

- *Surveillance* (e.g., using CCTV cameras at the perimeter line)
- *Guard* (e.g., located at a gate to the area)

If the perimeter is not well-protected by these security measures, both above mentioned threats will give rise to the threat *Crossing the grounds.* To mitigate this threat, two possibilities are given

- *Identity badges*
- *Patrols* (as described above)

If an attacker can overcome the area around the building, there are several ways to enter it. Thus, the threat *Crossing the grounds* gives rise to five new threats, i.e.,

- *Unauthorized entry into building (via door)*
- *Break-in door*
- *Break-in ground floor*
- *Break-in upper floor*
- *Break in lower ground floor*

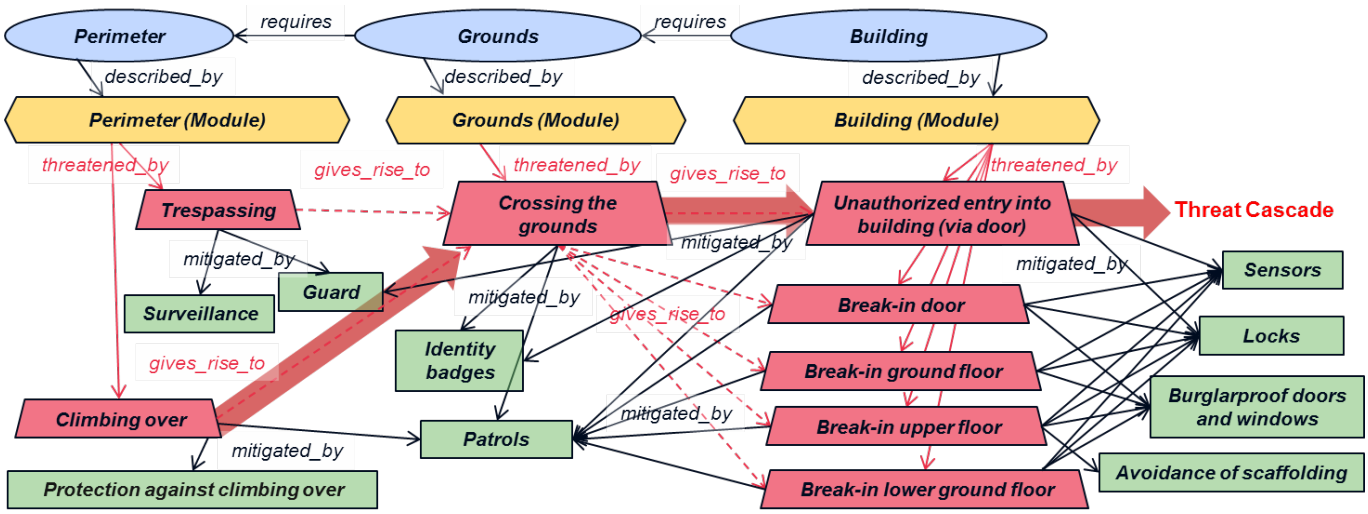

Figure 10. Graph-based model for the physical environment use case (excerpt).

For each of these threats, a number of safeguards are exemplarily given. One safeguard can be used to mitigate several of the above threats. For example, the safeguard Patrol is effective against all the threats. Other security measures could be

- *Sensors*
- *Locks*
- *Burglarproof doors and windows*
- *Avoidance of scaffolding*

Usually, the representation of the threatened objects (assets/modules, threats, safeguards) follows a risk inheritance model based on an instantiated structure similar to a fault tree. Such an implementation is based on master-scenario assets and the dependencies representation of the different branches of the attack tree is done as a nested set of a relational database. This approach described above is appropriate for the representation of different scenarios.

However, the physical scenario requires an individual definition and maintenance of the attack vectors, following separate paths for each variant. Consequently, such attack variants like trespassing/crossing the perimeter on the ground floor must be defined and maintained separately for the upper floors of the building.

To overcome this problem as well as to solve the basic requirement for a simple combination of risk information, the advantages of an implementation in a graph database (as described in Section V above), where nodes, relationships and attributes are provided, can be used. Additionally, the schema-less representation of a graph database simplifies extensions of the model. By applying graph operations, structures, traversals and aggregations can be easily extended as well. These advantages can only be achieved by switching to a representation within a graph database.

Based on the described threats within the physical security example and following the *_gives_rise_to_* relations of the ICT meta-risk model, the model will not only indicate one single but several potential threat cascades for the physical security use case (in Figure 10, we highlighted only one of them). Thus, due to the graph-based nature of the model, a large number of attack vectors can be described in a rather simple way. To compute a quantitative score for the risk analysis, the risk inheritance has to be modeled by appropriate functions (as already described in Section IV above).

VIII. RESULTS

In this section, it is demonstrated how the presented risk analysis approach can be used to derive (semi-)quantitative results (e.g., annualized loss expectancy (ALE) risk) based on semi-quantitative inputs (e.g., safeguard maturity levels according to the Capability Maturity Model Integration (CMMI) framework: 0.. Incomplete, 1.. Initial, 2.. Managed, etc.). By using graph databases as model environment, the writing of complex business code for risk estimation can be avoided by performing the required assessments using CYPHER statements.

The outlined risk estimation method is a simplified variant of the method defined in [20]. The general view is that vulnerabilities of assets can be exploited by threat sources resulting in negative impacts on protection criteria. Thus, for risk estimation, the vulnerabilities of assets are explicitly taken into account; however, instead of using them directly, they are substituted by maturity gaps of safeguards.

This results in risk as a function of the likelihood of the occurrence of a threat, the maturity gap of an associated safeguard, and the impact that the unwanted event has on protection criteria (cf. (1)).

$$R := f\left(T_{likelihood}, S_{maturity\ gap}, I_{protection\ criteria}\right) \quad (1)$$

In an initial step, the safeguard requirements are derived from goals and estimated using maturity levels (from [0…5]). The product of the maturity gap (i.e., 1+maturity gap to evade division by zero) and the safeguard priority (from [1…4]) gives an estimation of the *safeguard exposure* (from [1…24]) (2). Additionally, the relation to the potential maximum exposure (based on the current goal definitions) is also calculated (cf. (3) (4) and Figure 11).

safeguard exposure = *(1+ maturity goal – estimated maturity) * safeguard priority* (2)

safeguard exposure max = *(1+ maturity goal) * safeguard priority* (3)

safeguard exposure % = *safeguard exposure / safeguard exposure max*100* (4)

```
match (a:USE_CASE:Asset{name_de:'x'})
-[r1:described_by]->(b:USE_CASE:Module{name_de:'x'})
-[r2:threatened_by]->(c:USE_CASE:Threat)
-[r3:mitigated_by]->(d:USE_CASE:Safeguard)
with (1+r3.target_maturity-r3.maturity)*r3.priority as
  exposure, (1+r3.target_maturity)*r3.priority as
  exposure_max, r3
set r3.exposure = exposure
set r3.exposure_max = exposure_max
set r3.exposure_rel = r3.exposure/r3.exposure_max*100
return r3
```

Figure 11. Listing for the calculation of safeguard exposures.

After all safeguard exposures are calculated for each asset, the threat likelihoods are estimated for a specific timeframe (from [0…1], however, to simplify the CYPHER code, the null value is excluded to avoid a potential division by zero).

In a next step, the *threat exposures* are calculated. The threat exposure (from [0…20]) depends on the estimated likelihood (from [0…1]) and a function of its safeguard exposures (cf. (5)(6)(7) and Figure 12). For reasons of simplicity, here, the maximum function is used. In order to assess estimation variances, it may be appropriate to estimate the threat likelihood risk-averse (likelihood high) and risk-affine (likelihood low). Based on the calculation of current and potential maximum events, the risk factors within the model can be described either absolutely or relatively.

threat exposure = *likelihood(low) * MAX(safeguard exposure)* (5)

threat exposure max = *likelihood (high) * MAX (safeguard exposure max)* (6)

threat exposure % = *threat exposure / threat exposure max * 100* (7)

```
    match (a:USE_CASE:Asset{name_de:'x'})
    -[r1:described_by]-
>(b:USE_CASE:Module{name_de:'x'})-[r2:threatened_by]-
>(c:USE_CASE:Threat)
    -[r3:mitigated_by]->(d:USE_CASE:Safeguard)
with max(r3.exposure) as safeguard_exposure,
 max(r3.exposure_max) as safeguard_exposure_max, c
    set c.exposure = c.likelihood*safeguard_exposure
set c.exposure_max =
 c.likelihood*safeguard_exposure_max
    set c.exposure_rel = c.exposure/c.exposure_max*100
    return c
```

Figure 12. Listing for the calculation of threat exposures.

The threat exposures of all threats that have no incoming *gives_rise_to*-relationships are calculated first. The reason why the exposures of all uninfluenced threats are calculated initially is because no other threats have an effect on them (business impact analysis does not allow cyclic models).

After having calculated the threat exposures of all uninfluenced threats, the threat likelihood of all influenced threats (*gives_rise_to*-relations) can be updated based on the likelihood of their predecessors (chained likelihood). The calculation will be triggered as soon as all predecessors have been calculated. For reasons of simplicity, this is done by a simple multiplication of the original likelihood of the threat and the maximum of the likelihoods of its predecessors. Of course, a more complex function (weighting) representing the relative exposure of the threat to its influences can be used. In the following example (cf. Figure 13), the originally estimated likelihood of threat *'y'* is multiplied with the maximum of all its incoming *gives_rise_to*-likelihoods.

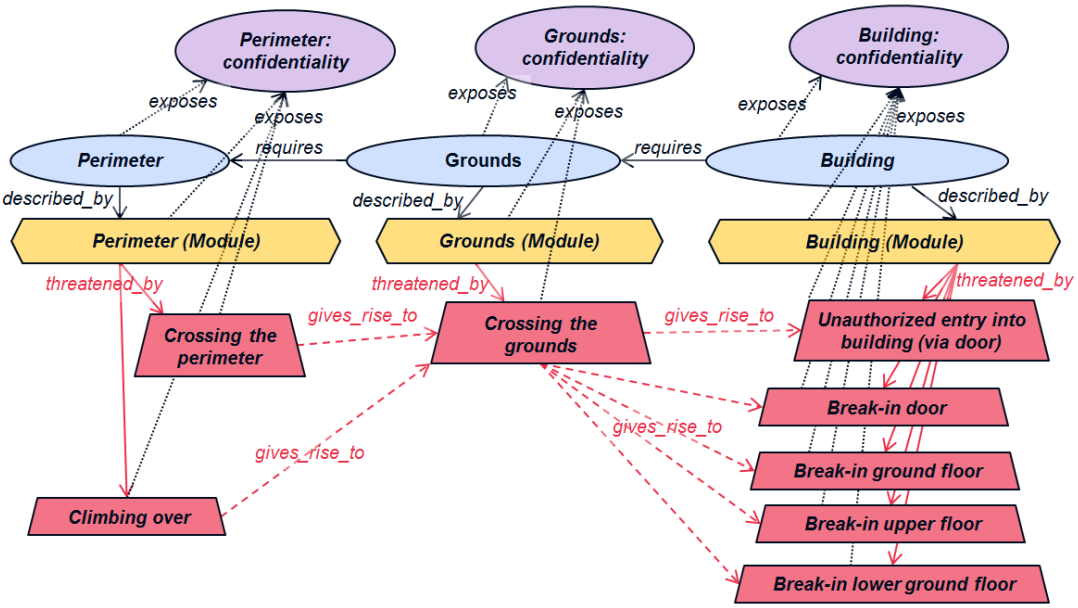

Figure 13. Possible aggregation of exposures on asset-specific protection criteria (here: confidentiality).

```
match (a:USE_CASE:Asset{name_de:'x'})
-[r1:described_by]->(b:USE_CASE:Module{name_de:'x'})
-[r2:threatened_by]->(c:USE_CASE:Threat)
-[r3:gives_rise_to]->(d:USE_CASE:Threat{name_de:
  'Threat y', module_id:2})
with max(c.likelihood) as trigger_likelihood,
  d.likelihood as original_likelihood, d
set d.original_likelihood = original_likelihood
set d. likelihood = d. likelihood *trigger_ likelihood
return d
```

Figure 14. Listing for the likelihood update of influenced threats.

After the likelihood update of all threats with incoming *gives_rise_to*-relations is finished, the remaining threat exposures can be calculated.

Depending on the desired level of detail, threats can be assessed individually or as generalized protection criteria related to assets (*asset exposures*), as illustrated in Figure 14. By extending the graph model, arbitrary aggregation layers can be defined. Here, to simplify the outlined use case, asset exposures are aggregated based on the maximum principle and risk is estimated based on the annualized loss expectancy (ALE) formula (cf. (8) and Figure 15). Again, additional lower and upper bounds could be integrated to express variance.

$$\textit{asset risk} = \textit{estimated impact} * \textit{MAX (threat exposure \% / 100)} \quad (8)$$

```
match (a:USE_CASE:Asset{name_de:'x'})
-[r1:described_by]->(b:USE_CASE:Module{name_de: 'x'})
-[r2:threatened_by]->(c:USE_CASE:Threat)
with max(c.exposure) as threat_exposure,
  max(c.exposure_max) as threat_exposure_max, a
set a.exposure = threat_exposure
set a.exposure_max = threat_exposure_max
set a.exposure_rel = a.exposure/a.exposure_max*100
set a.ale_risk = a.impact*a.exposure_rel/100
return a
```

Figure 15. Listing for the calculation of asset exposures and annualized loss expectancy (ALE) risks.

IX. CONCLUSION

This article describes how a generic ICT meta-risk model benefits from the qualities of a graph-based implementation, especially from the features of schema-less information, which can be parametrized based on the individual requirements of the organization, near-real-time traversals and flexible definitions of relationships between nodes, and the ability of easy model extension. A representative APT scenario is described to demonstrate a practical application of the presented ICT meta-risk model. The consideration of cascading risk effects, including human-based information system vulnerabilities, is a necessary prerequisite for an effective defense against APTs, which exploit the full range of attack vectors, from social over digital to physical. Consequently, a second use case introduced in this article illustrates an application of the ICT meta-risk model on a physical security attack. In general, the generic nature of the model allows addressing all kinds of threats - from the cyber over the physical to the business realm - and their dependencies.

The presented approach shows the application of a combination of several analysis steps and different parts of existing methods, e.g., morphological matrices, fault-tree- and event-tree-analysis, scenario analysis, threat analysis, system decomposition, and functional relationships [32]. The advantages of the presented combined approach are, for example, the possibility to focus on special requirements of information security and to cover a broader range of analysis depth and detail. These features cannot be achieved by using the previously mentioned methods on their own. The introduced scenarios are represented as a particular instance of a graph-based implementation of the generic ICT meta-risk model. The relevant risk components, which can be easily integrated into the graph-based ICT meta-risk model, are provided by widely-accepted ICT risk frameworks, most importantly by IT-Grundschutz. The defined relations between relevant risk components within this framework give an excellent starting point for possible paths that potential cascading risk effects might take.

From a technical point, for modeling and inference analysis of threat cascades, the graph-oriented database Neo4j with its query language CYPHER was used. Threat cascades and their relations can be visualized by graph databases in a more optimized way compared to relational databases. The schema-less data model of graph databases allows an easier adaption during the modeling process and the application of traversals to integrate calculations without modifications of the business code. However, with regard to the correctness of the results, the domain has to be specified and defined with a low level of uncertainty, and the level of detail of the risk factors has to correlate with the granularity of the results to guarantee a consistent distribution of risk values. Within the discussed use cases, uncertainty resulting from subjective assessments, or inconsistencies and errors in modeling depth is not dealt with explicitly. It can be addressed, like any other aspect, by introducing semi-quantitative descriptors (e.g., assessment uncertainty, etc.), which can be aggregated within the graph model similar to other variables.

ACKNOWLEDGEMENT

This work is party supported by the research project ”MetaRisk” (Project-Nr. 840905), which is funded by the Austrian National Security Research Program KIRAS (http://www.kiras.at/).

REFERENCES

[1] S. Schiebeck, M. Latzenhofer, B. Palensky, S. Schauer, G. Quirchmayr, T. Benesch, J. Göllner, C. Meurers, and I. Mayr, “Implementation of a Generic ICT Risk Model using Graph Databases,” presented at the SECURWARE 2015, 9th International Conference on Emerging Security Information, Systems and Technologies, Venice, Italy, 2015, pp. 146–153.

[2] T. W. Coleman, “Cybersecurity Threats Include Employees,” *International Policy Digest*. [Online]. Available: http://www.internationalpolicydigest.org/2014/05/12/cybersecurity-threats-include-employees/. [Accessed: 19-Mar-2015].

[3] SANS Institute, "Critical Security Controls: Guidelines." [Online]. Available: http://www.sans.org/critical-security-controls/guidelines. [Accessed: 19-Mar-2015].

[4] Ponemon Institute, "Exposing the Cybersecurity Cracks: A Global Perspective Part 2: Roadblocks, Refresh and Raising the Human Security IQ," Traverse City, Michigan, USA, 2014.

[5] Mandiant Intelligence Center, "APT1. Exposing One of China's Cyber Espionage Units," Mandiant, Alexandria, Washington, DC, Feb. 2013.

[6] D. Moon, H. Im, J. Lee, and J. Park, "MLDS: Multi-Layer Defense System for Preventing Advanced Persistent Threats," *Symmetry*, vol. 6, no. 4, pp. 997–1010, Dec. 2014.

[7] C. Tankard, "Advanced Persistent threats and how to monitor and deter them," *Netw. Secur.*, vol. 2011, no. 8, pp. 16–19, Aug. 2011.

[8] I. Friedberg, F. Skopik, G. Settanni, and R. Fiedler, "Combating advanced persistent threats: From network event correlation to incident detection," *Comput. Secur.*, vol. 48, pp. 35–57, Feb. 2015.

[9] The Commission on the Theft of American Intellectual Property, "The IP Commission Report," National Bureau of Asian Research, May 2013.

[10] G. Greenwald, *No Place to Hide: Edward Snowden, the NSA, and the U.S. Surveillance State*. Metropolitan Books, 2014.

[11] Internet Crime Complaint Center, "2013 Internet Crime Report," Federal Bureau of Investigation, 2013.

[12] BMI, "Polizeiliche Kriminalstatistik 2013," Bundesministerium des Innern, Berlin, 2013.

[13] International Organization for Standardization (ISO), Ed., *ISO 31000:2009 Risk management - Principles and guidelines*. ISO, Geneva, Switzerland, 2009.

[14] ISO International Organization of Standardization, *ISO/IEC 27005 - Information technology -- Security techniques -- Information security risk management*. 2011.

[15] BSI, "IT-Grundschutz-catalogues 13th version 2013," Bundesamt für Sicherheit in der Informationstechnik - Federal Office for Information Security, Bonn, Germany, 2013.

[16] International Standards Organization (ISO), *IEC 27004: 2009 Information Technology - Security Techniques - Information Security Management - Measurement*. Geneva, Switzerland: ISO, 2009.

[17] Federal Ministry for Transport, Innovation and Technology (BMVIT) and Austrian Research Promotion Agency (FFG), "KIRAS Security Research: MetaRisk," 2016. [Online]. Available: http://www.kiras.at/. [Accessed: 17-Feb-2016].

[18] T. Schaberreiter, "A Bayesian Network Based On-line Risk Prediction Framework for Interdependent Critical Infrastructures," Dissertation, University of Oulu, Oulu, Finlande, 2013.

[19] Chartis Research, "Looking fo Risk. Applying Graph Analytics to Risk Management. Leading practices from YarcData," 2013.

[20] S. Schiebeck, "An Approach to Continuous Information Security Risk Assessment focused on Security Measurements," Dissertation, University of Vienna, Wien, 2014.

[21] B. Williams and R. Hummelbrunner, *Systems concepts in action: a practitioner's toolkit*. Stanford University Press, 2010.

[22] S. Radeschütz, H. Schwarz, and F. Niedermann, "Business impact analysis—a framework for a comprehensive analysis and optimization of business processes," *Comput. Sci.-Res. Dev.*, vol. 30, no. 1, pp. 69–86, 2015.

[23] J. Göllner, T. Benesch, S. Schauer, K. Schuch, S. Schiebeck, G. Quirchmayr, M. Latzenhofer, and A. Peer, "Framework for a Generic Meta Organisational Model," in *Abstract Proceedings for the 14th FRAP Conference - Oxford*, Oxford, United Kingdom, 2014.

[24] R. McCrie, *Security operations management*. Butterworth-Heinemann, 2015.

[25] Neo4j Graph Database, "Intro to Cypher - Neo4j Graph Database." [Online]. Available: http://neo4j.com/developer/cypher-query-language/. [Accessed: 25-Mar-2015].

[26] R.-G. Urma and A. Mycroft, "Source-code queries with graph databases—with application to programming language usage and evolution," *Sci. Comput. Program.*, vol. 97, pp. 127–134, Jan. 2015.

[27] C. Batra and C. Tyagi, "Comparative Analysis of Relational And Graph Databases," *Int. J. Soft Comput. Eng.*, vol. 2, no. 2, pp. 509–512, May 2012.

[28] C. T. Have and L. J. Jensen, "Are graph databases ready for bioinformatics?," *Bioinformatics*, vol. 29, no. 24, pp. 3107–3108, Dec. 2013.

[29] ISACA, "COBIT 5 - Enabling Processes," Rolling Meadows, Illinois, 2012.

[30] T. UcedaVelez and M. M. Morana, *Risk Centric Threat Modeling: Process for Attack Simulation and Threat Analysis*. John Wiley & Sons, 2015.

[31] C. Wang and W. A. Wulf, "Towards a Framework for Security Measurement," in *Proc. of 20th National Information Systems Security Conference*, Baltimore, Maryland, 1997.

[32] Federal Aviation Administration (FAA), Ed., "FAA System Safety Handbook." 30-Dec-2000.