Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/cose

Computers & Security

Incorporating attacker capabilities in risk estimation and mitigation

Lotfi ben Othmane [a,*], Rohit Ranchal [b], Ruchith Fernando [b], Bharat Bhargava [b], Eric Bodden [a]

[a] Secure Software Engineering Group (SSE), Fraunhofer Institute for Secure Information Technology, Darmstadt, Germany

[b] Department of Computer Science, Purdue University, West Lafayette, IN 47907, USA

ARTICLE INFO

Article history:
Received 3 May 2014
Received in revised form
26 February 2015
Accepted 4 March 2015
Available online 25 March 2015

Keywords:
Risk estimation
Attack potential
Attacker capabilities
Risk mitigation
Threat
Uncertainty
Empirical research

ABSTRACT

The risk exposure of a given threat to an information system is a function of the likelihood of the threat and the severity of its impacts. Existing methods for estimating threat likelihood assume that the attacker is able to cause a given threat, that exploits existing vulnerabilities, if s/he has the required opportunities (e.g., sufficient attack time) and means (e.g., tools and skills), which is not true; often, s/he can perform an attack and cause the related threat only if s/he has the ability to access related resources (objects) of the system that allow to do so. This paper proposes a risk estimation method that incorporates attacker capabilities in estimating the likelihood of threats as conditions for using the means and opportunities, demonstrates the use of the proposed risk estimation method through two examples: video conferencing systems and connected vehicles, shows that changing attacker capabilities changes the risks of the threats, and compares the uncertainty of experts in evaluating the likelihood of threats considering and not considering attacker capabilities for two experiments. The results of the experiments suggest that experts are less uncertain about their estimations of threat likelihoods when they consider attacker capabilities.

1. Introduction

Developing a secure Information System (IS) requires assessing the risks to the IS and mitigating the identified threats. However, organizations have business goals and budget constraints that require addressing only a subset of the threats to the ISs they develop. Thus, they estimate the risk exposures of the threats (A function of the likelihood of the threat and the severity of its impacts (Wheeler, 2011)) and use the information to prioritize addressing the threats (McGraw, 2006).[1] The accuracy of risk exposure estimates leads to more realistic prioritization of the threats and therefore better return on investment.

Security experts produce widely different risk estimates because they have different opinions about the difficulty attackers have in exercising the threats against the system. The high uncertainty in the estimated risk exposure, indicated by

* Corresponding author.
E-mail addresses: lotfi.ben.othmane@sit.fraunhofer.de (L. ben Othmane), rranchal@purdue.edu (R. Ranchal), rfernand@purdue.edu (R. Fernando), bb@purdue.edu (B. Bhargava), eric.bodden@sit.fraunhofer.de (E. Bodden).

[1] Note that the goal of risk estimation is not to produce a set of numbers but to enable ordering the threats (Apostolakis, 2004).

http://dx.doi.org/10.1016/j.cose.2015.03.001

the high differences between the estimates, leads the business managers to view the exercise as of uncertain practical value and limits the ability of using the information to prioritize the threats (Boehm, 1991; Bonnette, July 2003; Wheeler, 2011). The main causes of the failure of security assessment according to Hubbard (Hubbard, 2009) are: failure to measure the effectiveness of the proposed method, use of methods that are found to include errors and biases, and do not use methods that are proven to work.

A commonly practiced approach in risk estimation is to identify all possible attack scenarios and estimate their risks (using maximum details). This approach is very costly and is not preferred in business projects. The alternative approach is to use a set of factors for estimating the risks of the threats grouped into classes using specific logic. (The methods we enumerate in this section fall in this category.) For example, the OCTAVE (Alberts and Dorofee, 2002) uses the factor classes: (1) motives to cause the threats; (2) means, which include required skills and knowledge to execute attacks and availability of tools; and (3) opportunities, which include time to perform the attack and number of allowed failed attempts. The factor classes used by NIST SP 800-30 (Stoneburner et al., 2002) are: (1) capabilities of the attacker, such as resources, expertise, and opportunities to perform attacks; and (2) intent of the attacker; that is, perseverance in attacking a specific asset to obtain sensitive information. (The definitions of the terms used in this paper are summarized in Table 1). Unfortunately, there is little formal guidance about the selection of the factors to use in the estimation models (Pardue et al., 2009).

Risk estimation methods, currently, separate the treatment of insider threats (violations of security policy by misusing granted privileges (Yasinsac, 2010).) and non-insider threats. However, in many cases, the same threat could be performed by insiders (entities that have access to data or resources (Bishop and Gates, 2008)) and non-insiders. For instance, an attacker who intends to cause the threat "interruption of a security camera of a corporation" and knows how to push the power off button of the camera, or knows how to cut the communication cable, cannot cause the threat unless s/he has the capability "physical access to the camera," where *attacker capability* is the ability to access a set of resources (objects) of the IS to exercise threats. Also, an attacker who plans to exercise the same threat and has time, expertise, knowledge, and tools to craft command messages to the camera to power it off cannot cause the threat unless s/he has the capability "inject messages to the network of the organization." Thus, attacker capabilities often conditions the use of acquired means and opportunities to cause threats (ben Othmane et al., 2013a).

Existing risk estimation methods commonly consider attacker capabilities as the resources (e.g., malware, scripts), knowledge, and expertise that could be used to cause threats, e.g., (Stoneburner et al., 2002). However, threats, often, could be exercised only if specific conditions are satisfied. These conditions include successful exercise of specific threats (e.g., getting insiders to collaborate), specific system configuration (e.g., VB script or ActiveX are enabled in the browser), and access to object resources that could be used in the attack scenarios. In this paper we investigate the use of attacker capabilities to access the resources (perform actions on the system resources) of the given system as conditions to use means and opportunities. (We consider attacker resources, knowledge and expertise as means to cause threats.) Previously, Duggan et al. considered accesses to resources (i.e., attacker capabilities) as a risk estimation factor (Duggan et al., Sep. 2007), but not as conditions for exercising the threats, which we do in this paper.

This paper discusses the use of attacker capabilities in estimating the likelihoods of threats and shows how considering attacker capabilities, as extra information given to the

Table 1 – Definitions of used risk-related terms.

Term	Definition
Access	A specific type of interaction between a subject and an object that results in the flow of information from one to the other (National Computer Security Center (NCSC), 1988).
Asset	Things that have values and are required to achieve the goals of the IS (Dubois et al., 2010).
Attacker capability	The ability to access a set of resources (objects) of the IS to exercise threats.
Impact	The potential consequence of a risk that may harm the assets of a system or an organization (Dubois et al., 2010). It could be financial, legal, operational, damage reputation, and privacy violation.
Means	The tools, skills, and knowledge required to perform actions that cause the given threat. (cf. (Alberts and Dorofee, 2002)).
Resource	An entity (Object) that contains or receives information, such as records, files, programs, video displays, and devices (National Computer Security Center (NCSC), 1988). We use the terms object and resource in this paper interchangeably.
Opportunities	The circumstances that make attacking the system possible, such as the time to perform the attack and the number of allowed failed attempts (cf. (Alberts and Dorofee, 2002)).
Risk	The combination of a threat with one or more vulnerabilities leading to an impact harming one or more of the assets (Dubois et al., 2010).
Risk exposure	A function of the likelihood of the threat and the severity of its impacts (Wheeler, 2011) (Bohem used close terms to define risk exposure (Boehm, 1991)).
Security policy	A statement of what is, and what is not, allowed (Bishop, 2012).
Threat	Potential attacks, carried out by agents, that target ISs's assets (Dubois et al., 2010). In general, it is a potential violation of security policies of the given system (Bishop, 2012).
Threat agent	An agent that can cause harm to assets of the ISs (Dubois et al., 2010).
Threat likelihood	Measures the frequency and possibility that the given threat occurs (cf. (Wheeler, 2011)).
Threat severity	Measures the impacts of the given threat in terms of losses and damages (cf. (Wheeler, 2011)).
Vulnerability	A characteristic of an IS asset or group of IS assets that can constitute a weakness or a flaw in terms of IS security (Dubois et al., 2010).

experts in estimating the risks of threats (ben Othmane et al., 2013a), improves the accuracy of risk estimation. The main contributions of the paper are: (1) propose a method for incorporating attacker capabilities in estimating the likelihood of threats as conditions for causing the threats; (2) illustrate the method through two examples; (3) demonstrate that removing attacker capabilities can reduce security risk to ISs; and (4) evaluate the effect of considering attacker capabilities on the uncertainty in estimating the likelihoods of threats through 2 experiments. This work shows that the information "attacker capability," which we propose to use, improves the quality of the risk estimates (since it reduces the uncertainty in the estimates) in the later approach.

The paper is organized as follows. First, we give an overview of related work (Section 2). Then, we discuss attacker capabilities and their likelihoods (Section 3), we propose a risk estimation method that incorporates attacker capabilities (Section 4) and discuss how to mitigate risks through changing attacker capabilities (Section 5). Next, we illustrate the use of the method through two examples: a video conferencing system and connected vehicles (Section 6), evaluate the proposed risk estimation method (Section 7), and discuss the impacts of the results on the state of the art (Section 8). We conclude the paper afterwards.

2. Related work

We describe in the following a set of publications that describe risk estimation methods close to the work we present in this paper and we show how our work is different.

The TREsPASS Project (The TREsPASS Project, Oct. 2014) aims to develop tools to systematically support, predict, prioritize, and prevent complex attacks that cleverly exploit multiple vulnerabilities, involving physical infrastructures and human. The expected results of the project should help the defenders to make rapid decisions regarding which attacks to block since, for example, the attackers can gain knowledge very fast and the infrastructure can change rapidly.

The Open Web Application Security Project (OWASP) (OWASP, December 2011) proposes an approach for estimating security risk that combines the likelihood and severity of each threat using inference rules. Likelihood of a threat is estimated through evaluating a set of factors that measure the threat agent (the entity that causes the threat, e.g., attacker) and chances to discover and exploit vulnerabilities related to the threat. The severity of a threat is estimated by evaluating a set of factors that measure its business and technical impacts. Business impacts include financial losses and privacy violations. Technical impacts include loss of confidentiality, integrity, availability, or accountability. The scores of the factors of the likelihood of a particular threat are added up and divided by the number of factors. The resulting score is transformed using a function that maps scores to qualitative values—e.g., low, medium, and high. The same applies for severity.

Ekelhart et al. (Ekelhart et al., 2009) developed a prototype of a framework for information security risk management, AURUM, for supporting NIST 800-30 risk management process (Stoneburner et al., 2002). The framework uses a security ontology to describe the IS, a threat catalog based on the German IT Grundschutz (BSI, August 2012) and the EBIOS (ANSSI, Jan. 2010) threat catalogs, a vulnerabilities catalog, and a catalog for security controls (countermeasures that include policies, processes, procedures, organizational structures and software and hardware components (ISO/IEC 27002, 2005).) for ISs' resources/assets. The framework determines the likelihood of a given threat using the likelihood of the threat in previous estimates, the likelihood of threats that cause the given threat, and the likelihood of exploiting related vulnerabilities. In addition, the framework recommends controls that mitigate identified risks, evaluates the security controls, and analyzes their costs and benefits. It also includes an interactive tool that allows the decision makers to evaluate security solution strategies given a set of costs and benefits criteria, such as revenue, reputation, and performance.

Chivers et al. (Chivers et al., 2009) observe that most threat assessment methods are based on identifying threat paths used by attackers to compromise assets, which link point of entries of the given system to the assets they target. These methods ignore the reality that the attackers compromise the system components that allow them to reach the assets of the system they are trying to compromise. They propose the use of risk profile, such that each component of the system is associated with a risk profile: a set of attacks that, after compromising the system's components, enables the attackers to attack the system. Basically, attackers exploit the trust (and privileges) between the components of the given system to reach from an initial component that they compromise to other system components that manage the targeted assets. The proposed approach supports the propagation of security risk among the components of the system and enables readjustment of the security risk to the system when components change—without the need for a full reevaluation of the risk of the new system.

Wheeler (Wheeler, 2011) proposes a qualitative model (i.e., uses fuzzy values like low, high) that measures the security risk considering the sensitivity of the resources, the likelihood of the threats, and the severity of exploiting the vulnerabilities. The sensitivity aspect measures the consequence of the threat for the organization, e.g., financial loss. The severity aspect measures the magnitude of the vulnerability independently of the threat source and the asset sensitivity in terms of affected confidentiality, integrity, availability, and accountability. The likelihood of a threat measures the probability that the related vulnerabilities are successfully exploited and the frequency of the occurrence of the threat. The main factors used to measure threat likelihood are: the size of the threat universe (the scope of the user community that can exploit the vulnerability, e.g., Internet users, insiders, or administrators), motivation of threat actors, sophistication of the given attack or required level of skills, knowledge of organization and system, level of existing security controls to mitigate the threat, and attractiveness of the target. The model uses relative scales to rate risk based on predefined criteria for each level of the scale and relies on the knowledge and experience of the assessor to apply the scale to the threats.

The approach that we propose considers that threat likelihood measures the likelihood to exploit vulnerabilities

related to the given threat and the frequency of the occurrence of the threat, which is similar to the approach proposed by Wheeler (Wheeler, 2011); relies on the knowledge and experience of risk assessors to use a set of factors and scales to estimate the risk of the evaluated threats, which is similar to the approaches proposed by Wheeler (Wheeler, 2011) and OWASP (OWASP, 2011), and supports the propagation of security risk among the threats to the system as proposed by Chivers et al. (Chivers et al., 2009) and Ekelhart et al. (Ekelhart et al., 2009).

3. Attacker capabilities and their likelihoods

This section defines attacker capabilities and describes how to identify them and to estimate their likelihoods.

3.1. *Description of attacker capabilities*

Macmillan Dictionary defines capability as the "ability to do something" and the "number of weapons, soldiers, etc. that a country has for fighting a war" (Dictionary, 2014).[2] This definition is aligned with the current use of term attacker capability, such as in NIST SP 800-30 (Stoneburner et al., 2002); that is, the means that a given attacker has—i.e., skills, knowledge, time, expertise and tools.

A cyber security attacker could have the motives, means, and opportunities (i.e., OCTAVE classes of factors) to cause a specific threat or the capabilities and intents (the NIST classes of factors). However, s/he *can* often perform an attack and cause the related threat only if s/he has the capability to access or use a related system resources (objects) to do so. For instance, an attacker who wants to eavesdrop on messages exchanged between two parties and has the required means and opportunities to cause the threat cannot do so *unless* s/he can "access the communication link" between the two parties or to one of the terminals used in the communication. We call this ability to access a set of resources (objects or components of the system including communication links between the components) of the IS to exercise threats an *attacker capability*.[3] Attacker capabilities are conditions for causing the threats to a given IS, not as information describing the potential attackers. This condition does not apply to all threats. For instance it does not apply to most social engineering threats.

Security policies specified by the management (management access policies) could be formulated using the tuple <*S*,*O*,*A*,*E*>, where *S* is the set of subjects, *O* is the set of objects, *A* is the set of actions, and *E* is the set of conditions, including purposes and contexts of the access (Bishop et al., 2010)—The objects of the system are the system resources[4] and actions (or operations (Hu et al., Sep. 2006)) are the access types (or access methods) on these objects. Ideally, these policies should be the source of attacker capabilities. However, implementing access policies results, sometimes, in providing the subjects with more privileges than the ones specified by the policies (Bishop et al., 2010). The reasons for the differences between the management access policies and the *effective access policies* include the difficulty to implement access conditions, such as intents. For instance, insiders could misuse their access privileges to exercise threats, e.g., an insider who has physical access to the server hosting an online Web application (e.g., online booking of airline tickets) and aims to interrupt the application could perform the attack by (just) unplugging the power cable of the server. Thus, the effective access policies are the source of attacker capabilities, which could be formulated using the tuple <*S*,*O*,*A*>.[5]

The concept of "attacker capability" extends the Information Systems Security Risk Management (ISSRM) domain model proposed by Dubois et al. (Dubois et al., 2010). Fig. 1 depicts the conceptual model of security risk assessment considering attacker capabilities. The definitions of the terms used in this model are provided in Table 1 [6] In this model, a threat, when it occurs (i.e., as an event), compromises an asset and has a set of impacts. A threat could be caused by exploiting a vulnerability, e.g., an attacker intercepts messages exchanged between two parties and exploits the weakness that data are in plain-text. (Further discussion about the ISSRM domain model and how it is developed is in (Dubois et al., 2010).)

The shaded area of Fig. 1 includes the concepts that extend the model: "motive," "means," "opportunity," and "capability." The concepts "motive," "means," and "opportunity" are commonly used in the literature with slightly different meaning, e.g., in (Alberts and Dorofee, 2002; Risk Steering Committee, Sep. 2008). The concept "capability" is the new concept that we introduce in this paper. A threat agent requires means, opportunities, and capabilities to perform attacks. Means include skills, knowledge, and tools; opportunities include time to perform a successful attack; and capability is the ability to access or use a set of system resources to exercise threats; that is, the ability to use appropriate means and opportunities to cause the threat by exploiting related vulnerabilities. The means that a given threat agent *could use* to cause a given threat depend on the capabilities that s/he has.

Note that the motives to cause a given threat may not depend on attacker capabilities. For instance, an attacker frustrated by the limited flow of water in the pipe that supplies his/her house may *aim* to stop the mechanism (which is the goal) through e.g., changing the firmware of the device that limits the water flow (which is the threat). Some threats could also be exercised without the need for attacker capabilities. This, for example, applies to social engineering threats. (Social

[2] These definitions are aligned with the ones of the Oxford dictionaries and the Cambridge dictionary.

[3] Attacker capabilities are with respect to ISs.

[4] It is common in the access control literature to use the term resources for the objects of the given system, e.g., (Bishop et al., 2010) and (Hu et al., Sep. 2006).

[5] Bishop et al. have previously used the tuple in formulating "feasible" policy (Bishop et al., 2010).

[6] There are variety of definitions in the literature for each risk-related concept/term. We use, when possible, the definitions provided by Dubois et al. (Dubois et al., 2010), since the work aligns the different definitions—which helps to have a common understanding of the concepts.

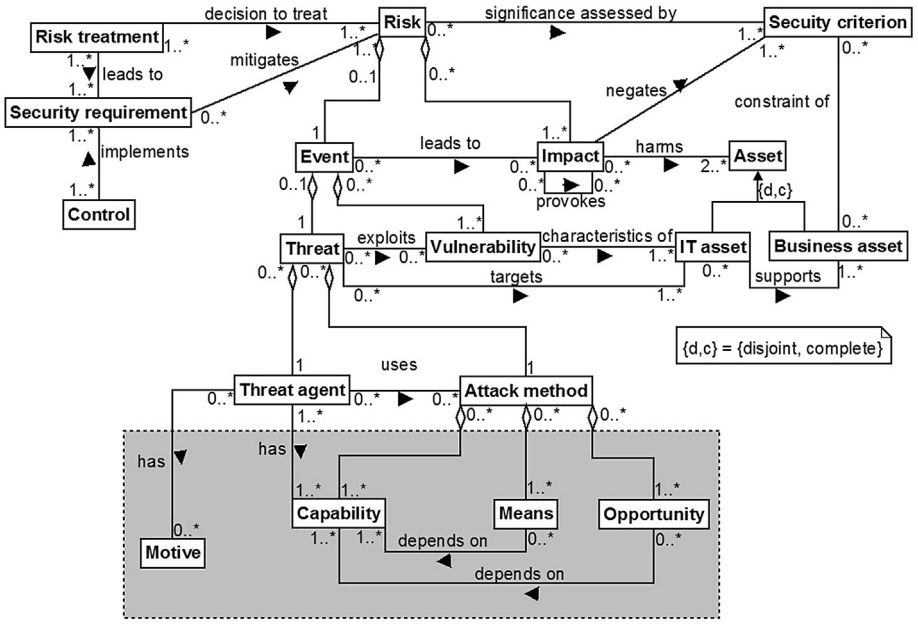

Fig. 1 – Conceptual model of the risks of security threats. (The concepts in the shaded areas extend the ISSRM domain model. (Dubois et al., 2010)).

engineering threats are not attacker capabilities because they are not accesses to the resources).

3.2. *Identification of attacker capabilities*

The identification of attacker capabilities for a given IS requires identifying all the access modes to the given system. The attacker capabilities could be formulated using the tuple <O,A>, where O is the set of objects and A is the set of actions; we discard the subjects from the tuple because we consider that all are potential attackers. The elements of the tuple could be derived from the system architecture[7] as follows:

1. Decompose the system into objects/components;
2. Identify the type of each object, e.g., hardware, software, database and communication link;
3. Enumerate all possible accesses to each object. This could be performed using a template that lists the attacker capabilities for each object type that attackers need when using the given object to perform attacks—see e.g., Table 2.

Note that we use communication link to refer to data in transit on a communication link in Table 2. (It is also possible to use data objects, not communication link.) This gives the attacker perceptive view of the accesses to the resources, e.g., view "read messages" as "intercept messages."

We claim that these steps are sufficient to allow the process to identify all the (known) attacker capabilities of the system. First, the use of the system architecture assures the enumeration of all the objects/components that compose the given system. Second, the use of attacker capabilities template assures enumeration of all (known) attacker capabilities for the given object.[8]

3.3. *Estimation of capability likelihood*

Attacker capabilities likelihood measures how likely potential attackers could have the given capability, which allows them to attack the ISs. The metric could be computed using historical data if it were available.

An alternative approach is to have experts who estimate the chances that attackers could have the capabilities. E.g., each expert (or a set of experts) assigns a score that measures the expectation that attackers could have the given capability c using the scale 0 to 4, where 4 means that certainly the attackers could have the capability, 3 indicates the attackers may have the capability, 2 indicates the attackers could have the capability, 1 indicates the attackers are unlikely to have the capability, and 0 indicates that it is very unlikely (still may be possible) they have the capability. To have capability likelihood values in the range 0–1, we could map the scores to values in the interval 0 to 1. The experts could use a set of factors to estimate the attacker capabilities, which depend on the operational context of the given system. These may include the location of the system resources—e.g., cameras in the streets, the number of the instances of the specific resource (e.g., number of sensors available for use), the number of potential attackers that can have the capabilities, the distance that the attackers have to travel to perform to access the resource, i.e., the distance between the potential attackers and the given resource.

[7] The architecture of an IS is the set of structures that compose it, including the software and hardware elements, relations among them, and properties of both (cf. (Clements et al., 2011)), where the (object) resources are the system components.

[8] The enumeration approach is similar to the one used by the STRIDE approach for threat modeling (Howard and Lipner, 2006).

Table 2 – Example of attacker capabilities template.

Object type	Example of capabilities
Hardware device	Physical access.
Software component	Local access (e.g., copying).
	Communicate with the component from within the local domain.
	Communicate with the component from outside the local domain.
Database	Query the database and
	Modify the data files.
Communication link	Intercept messages and
	Modify messages
	Remove messages.

For some attacks a potential attacker needs a set of l capabilities c_k that are required *all together* to perform threat t. We formulate this case using Equation (1), which computes the likelihood of capability C from the likelihoods of the l capabilities.

$$C = \prod_{k=1}^{k=l} (c_k) \tag{1}$$

4. Risk estimation method considering attacker capabilities

This section proposes an approach for estimating the likelihood, the severity, and the risk of security threats. It uses a set of symbols described in Table 3.

4.1. Overview of the risk assessment process

Risk assessment determines the risks associated with IS (Stoneburner et al., 2002). The process for assessing security risks includes identifying the threats to the system, estimating their risks and identifying controls for mitigating the risks—i.e., addressing the threats. The main steps of the process, as shown in Fig. 2 (cf. (Stoneburner et al., 2002)), are:

- Describe the system. We identify the boundaries[9] of the IS and its software architecture; that is, the components and the relationships among them and the environment.
- Identify the threats. We identify the assets[10] to protect and the motivations, and capabilities (using approach described in Subsection 3.2) of the attackers. Then, we identify the threats applicable to the IS. We can use as guidelines the German IT Grundschutz threat catalog (BSI, August 2012), the EBIOS threat catalog (ANSSI, Jan. 2010) and the STRIDE taxonomy (Hernan et al., November 2006).
- Identify the vulnerabilities. We identify flaws and weaknesses of the system that could be exploited by attackers.

[9] A domain within which a particular security policy or security architecture applies.

[10] A system resource required to be protected by the information system's security policy, intended to be protected by a countermeasure, or required for a system's mission (Shirey, Aug 2007).

Table 3 – List of symbols used in the method.

Sym-bol	Description
$R(t)$	Risk estimate of threat t.
m	Number of factors used to estimate the severity of the impacts of threats.
n	Number of factors used to estimate the ease of causing the threats.
$Fi_j(t)$	Evaluation, in terms of score, of the impact factor j of threat t.
$Fl_j^c(t)$	Evaluation, in terms of score, of the likelihood factor j of threat t with respect to capability c.
$C_k(t)$	Likelihood that potential attackers have capability k.
$S_t^{c_k}$	Likelihood of easiness of causing threat t using capability c_k.
$S(t)$	Likelihood of easiness of causing threat t.
$O(t)$	Likelihood of occurrence frequency of threat t.
$L(t)$	Likelihood of threat t.
$I(t)$	Likelihood of the impact of threat t.

- Analyze the security controls. We analyze the security controls that are implemented or are planned for implementation to reduce the likelihood or severity of threats. A control is either technical or nontechnical. Technical controls are safeguards incorporated into hardware or software, such as access control mechanisms and intrusion detection systems. Nontechnical controls are management and operational controls, such as security policies and operational procedures.
- Estimate the likelihoods of threats. We derive scores that indicate the expectation that the threats occur considering potential vulnerabilities of the system, the effectiveness of existing security controls, the attacker capabilities and motivations, and the means and opportunities that potential attackers may use. We describe the estimation process in more details in Subsection 4.2.
- Estimate the severities of the threats. We identify the consequences of causing threats to the system, which can be financial, legal, operational, safety, privacy, or reputation damage. We describe the estimation process in more details in Subsection 4.3.
- Estimate the risk. We compute the risk exposure of each threat to the system. We describe in more details the process in Subsection 4.4.
- Recommend security controls. We identify security controls that mitigate the identified threats or reduce the risk to an acceptable level.

4.2. Estimating threat likelihood

We define the *likelihood* of a threat as the expectation that the threat occurs. The objective approach to estimate the likelihood of a given threat is to collect historical data (e.g., events log of the operating systems) and use them to compute the frequency of occurrence of the threat during a specific period. However, data are rarely available and we have to rely on expert opinions to estimate the likelihood of threats.

The likelihood of a given threat combines two scores: ease of causing the threat and threat occurrence frequency. The *ease of causing the threat* t estimates the difficulty attackers have

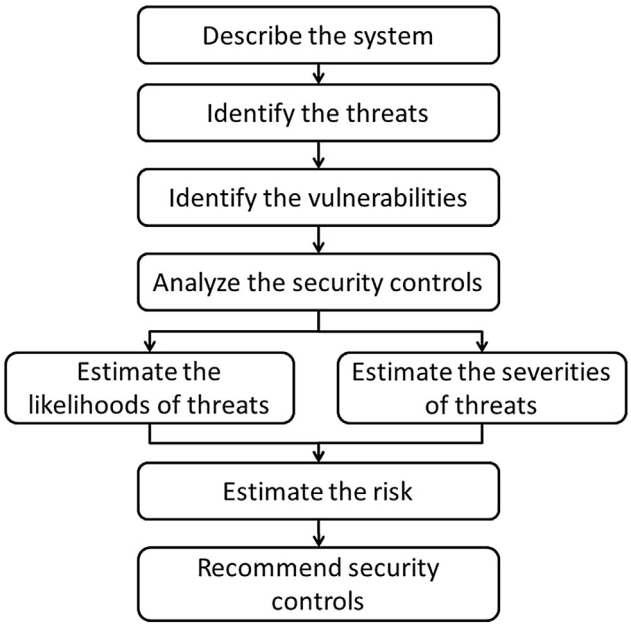

Fig. 2 – Risk assessment process.

to attack the system and cause the threat, which is measured as the expectation that potential attackers have the required capabilities, opportunity and means to cause the given threat. The *threat occurrence frequency* measures the expectation of the frequency that attackers cause the threat, assuming they have required capability, means, and opportunities. In the following, we discuss how to estimate both quantities.

4.2.1. *Estimating ease of causing a threat*

The experts assign scores to factors[11]: elapsed time, specialist expertise, knowledge of the system, window of opportunity, and required equipment and tools; they select for each factor the level of complexity/difficulty required to develop attacks that cause the evaluated threat. The levels of complexity/difficulty are mapped to numerical values based on predefined scales such as the ones in Table 4.

Let C_k be a capability required to cause threat t and assume we have the evaluations of the n factors $\{Fl_1(t),\ldots,Fl_n(t)\}$ in terms of scores provided by a technical expert. The score of ease of causing threat t is the sum of the scores of the factors multiplied by the capability likelihood. Equation (2) formulates the computation of ease of causing threat t considering attacker capability C_k.

$$S_t^{C_k} = \left(\sum_{j=1}^{j=n} Fl_j^{C_k}(t)\right) \times C_k(t) \tag{2}$$

Some threats could be associated with several options of capabilities. In this case, the score of ease of causing a threat is the maximum of the scores of ease of causing the threat using all possible capabilities. Equation (3) formulates the computation of the ease of causing threat t considering all attacker capabilities.

$$S(t) = max\{S_t^{C_k}\} \tag{3}$$

Recall that there are several other factors that condition causing threats, such as system configurations—as discussed in the introduction. The equations we provide above do not consider these conditions, however they could be extended to do so.

4.2.2. *Estimating threat occurrence frequency*

We define the likelihood of occurrence frequency of threats as the likelihood of attacker motivations—e.g., financial gain, fun, fame and technical advancement. Business experts (who are familiar with the business aspect of the system) identify attacker motivations in attacking the system and use the information to provide an estimate of the threat occurrence frequency.[12]

4.2.3. *Estimating the likelihood of a threat*

Equation (4) provides the formula for computing the likelihood of threat t; which combines the score of ease of causing the threat, $S(t)$, and the score of the threat occurrence frequency, $O(t)$.

$$L(t) = S(t) \times O(t) \tag{4}$$

4.3. *Estimating threat severity*

The experts associate scores for the factors: financial, legal, operational, safety, privacy and reputation damage (ISO/IEC, 18045, 2008), which are described in Table 5. They select for each factor the level of magnitude of losses or harms that the evaluated threat causes. The magnitudes of losses and harms are mapped to numerical values based on predefined scales such as the ones in Table 5.

Assume we have the evaluations of the m factors $\{Fi_1(t),\ldots,Fi_m(t)\}$ in terms of scores provided by a business expert. The score of severity of threat t is the sum of the scores of the factors. Equation (5) formulates the computation of the severity of a given threat t.

$$I(t) = \sum_{j=1}^{j=m} Fi_j(t) \tag{5}$$

Note that business experts could use the impact factors of Table 5 or other impact factors that better describe the impacts of the threats for the IS being assessed.

4.4. *Estimating the risk exposures of threats*

Equation (6) formulates the risk exposure of threat t, which combines the likelihood of the threat and its severity. The risk exposure scores are mapped to risk exposure levels in the scale 0 to 1 using a function such that the possible maximum score is mapped to 1 and 0 is reserved for score 0.

[11] The factors are close to the ones used in ISO18045 (ISO/IEC, 18045, 2008) to estimate the efforts required to create and demonstrate an attack that exploits a vulnerability.

[12] Computing the occurrence frequency of a threat as the sum of the attacker motivations for causing the threat is an overestimate because the motivations are not totally independent.

Table 4 – Factors for estimating the ease of causing the threats.

Factor	Description	Example of scales
Elapsed time	Time taken to identify a vulnerability related to the threat being analyzed, and to develop and perform an attack that causes the threat.	2 for 1 year, 4 for 1 month, 6 for 1 day and 8 for few minutes.
Specialist expertise	Level of generic knowledge about the product type—e.g., protocols, operating system, algorithms—and the principles required to cause the threat.	8 for layman, 6 for professional, 3 for expert and 1 for multiple experts.
Knowledge of the system	Specific expertise about the system—e.g., configuration parameters, location of files, etc.	2 for deep knowledge is required, 4 for generic knowledge is required and 8 for no knowledge is required.
Window of opportunity	Number of samples that the attacker can obtain or number of attacks without identification—e.g., number of password trials before the system disables the username.	8 for unlimited, 6 for one year, 4 for one month, 2 for one day and 1 for few minutes.
Required equipment and tools	Equipment and tools required to identify and exploit vulnerabilities related to the given threat, e.g., equipment to perform power analysis of a smart card (Messerges et al., 2002).	0 for not available, 2 for available only to experts, 4 for expensive equipment and 8 for cheap equipment or script available on the Internet.

$$R(t) = I(t) \times L(t) \quad (6)$$

5. Risk mitigation through changing attacker capabilities

Software will be often shipped with known and future vulnerabilities and many of these vulnerabilities will be discovered and exploited. We can however reduce the risk exposure associated with the exploitation (Manadhata and Wing, 2011). We propose, in the following, reducing risk through changing attacker capabilities and a metric to measure the efficiency of changing attacker capabilities.

5.1. *Reducing security risk through changing attacker capabilities*

Recall that attacker capabilities are conditions for exercising threats; that is, an attacker *can* only cause a threat *t* if s/he has one of the required capabilities for the threat. Therefore, one approach to mitigate a given threat is to deny potential attackers the required capabilities, or the ones with high likelihoods. Thus, a change in attacker capabilities may substitute the need to implement a set of security controls.

The techniques to deny attacker capabilities to IS depend on the capabilities themselves. Nevertheless, there are two common techniques that commonly apply. The first technique is changing the IS architecture; that is, changing the software and hardware components of the IS and/or relations among them. An example follows. Let an ISs be composed of devices given to customers that communicate with a remote Web application. Assume the system requires the use of a secret computation that could be deployed either in the devices or in the Web application. An attacker who wants to learn the secret computation needs either to get the code from one of the devices or get it from the remote office. An approach to mitigate the threat is to deploy the sensitive computation in the remote office because it is highly likely that attackers can get access to a device, extract the code,

Table 5 – Factors for estimating severity of threat impact.

Factor	Description	Example of scales
Financial	Losses of revenue and cost for repairing the IS. Losses are measured in terms of monetary value, e.g.–above 100 000 €.	0 for no loss, 1 for low, 2 for moderate, 3 for heavy and 4 for very heavy loss.
Legal	Legal impacts resulting from the threat, such as intellectual property theft, potential lawsuit from customers.	0 for no legal issues, 1 for customers do not notice the legal infractions, 2 for customers do notice the legal infractions, 3 for legal complaints and 4 for penalty due to legal complaints.
Operational	Duration of interruption of services of some or all the components of the IS, e.g., interruption of gateway, BackOffice (BO) for 1 minute.	0 for no interruption, 1 for interruption but customers are not aware, 2 for interruption and customers are aware, 3 for system down for few hours and 4 for system down for few days.
Safety	Harms resulting from the threat and their level, e.g., light injuries for one person and death of several persons.	0 for no injuries, 1 for light injuries, 2 for severe injuries, 3 for few fatal injuries, 4 for multiple fatal injuries.
Privacy	Losses of private information of customers or employees, e.g., unauthorized disclosure of sensitive data of customers to third parties for second use without their agreements.	0 for no data access, 1 for anonymous access, 2 for specific data for a set of customers, 3 for identity compromise of a customer and 4 for identities of many customers are compromised.
Reputation damage	Damage to the reputation of the IS and the organizations, e.g., customers' dissatisfaction, project termination, organization closure.	0 for no damage, 1 for customers are not happy, 2 for losing some customers, 3 for losing the big customers and 4 for decommissioning the system.

reverse engineer it, and get the secret computation; that is, potential attacker are highly likely to be able to access a device of a customer that runs the code.

The second technique of denying attacker capabilities is to change access control policy of the IS; that is, to change the access privileges of the users on the resources of the system. An example follows. Assume an attacker, who is an employee of an organization, has the capability to remotely access the server that hosts the components of the IS, which allows him to interrupt the server. An approach to mitigate the threat is to change the policy such that only system administrators can access the server. So, a change of the policy could deny potential attacker capabilities and avoid related threats.

Note that in some cases it is not possible to deny the attacker capability.

5.2. *Evaluating the efficiency of changing an attacker capability*

Assume we have an IS which has two threats T_m and T_k, where threat T_m depends on capability C_m and threat T_k depends on capability C_k. Let R_m^O be the risk of threat T_m and R_k^O be the risk of threat T_k. Assume we change the architecture of the IS such that attackers acquire capabilities C_h and C_l and lose capability C_m and C_k. Let R_m^N be the new risk of threat T_m, which requires capability C_h, and R_k^N be the new risk of threat T_k, which requires capability C_l. We consider that an IS architecture change reduces the risk of the system *if* the sum of the risks of the threats after applying the change is smaller than the sum of the risks of the threats before applying the change; that is, the change improves the security of the IS if and only if Equation (7) holds true.

$$R_k^N + R_m^N - \left(R_k^O + R_m^O\right) < 0 \tag{7}$$

We define *efficiency of capabilities change* as the relative risk improvements resulting from the capability changes. Efficiency metric is useful in the case of choosing a software architecture option among a set of alternatives. Equation (8) formulates the efficiency of such operation as the risk improvement divided by the initial risk (i.e., risk level before the change activity) of the system. The results of the equation could be used to rank alternative options.

$$E_c = \frac{R_k^N + R_m^N - R_k^O - R_m^O}{R_k^O + R_m^O} \tag{8}$$

6. Examples of using capability-based risk estimation and mitigation

This section illustrates the use of attacker capabilities in estimating security risks of two example systems: video conferencing systems (Subsection 6.1) and connected vehicles (Subsection 6.2).

6.1. *Example 1: video conference system*

This section illustrates the use of our proposed approach using as example a Video Conferencing (VC) system. First, we describe the system including the assets. Then, we estimate the risks associated to the system. Next, we give examples on how changing attacker capabilities reduces the risks to the system.

6.1.1. *Description of the system*

A VC system, as the one illustrated by Fig. 3, allows two or more users, each equipped with a VC station, to communicate and share applications and documents, while being at different locations. It is commonly used, for example, in business meetings and online lectures. The parties can (1) capture and exchange video, audio, and data, and (2) control the devices of the remote station, e.g., change the direction of the camera or zoom on an object. The devices at a station include: camera for video input, television for video output, microphone for audio input, speaker for audio output, and computer for data and application sharing.

Motivated malicious attackers could be interested to snoop on organization meetings stealthily, view documents in the meeting rooms, or disturb the meeting schedules of the staff of the organization and their partners (cf. (Perlroth, January 2012)). They cause threats to the system that allow them to achieve their goals.

In the following we illustrate the use of the proposed method for estimating the risks of threats to the VC system

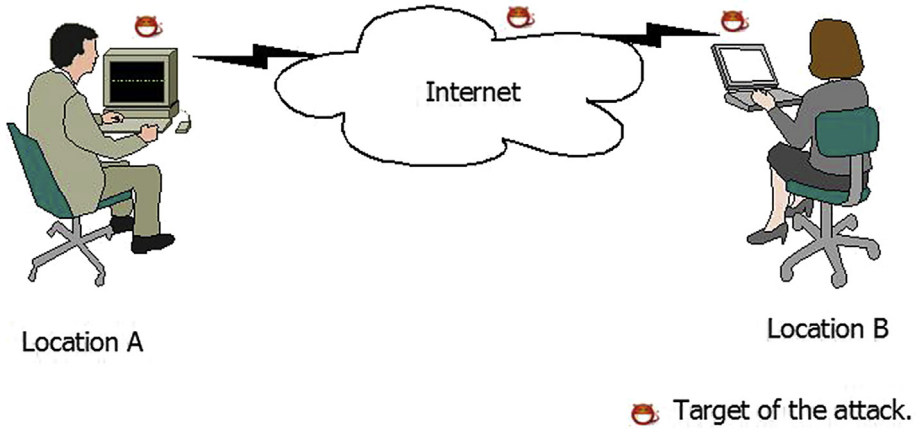

Fig. 3 – Video Conferencing System and associated attacks.

Table 6 – Sample of attacker capabilities corresponding to the system.

ID	Capability	Score	Likelihood
C1	The attacker can discover the VC station (e.g., through port scan.) and the station has automatic reply feature on.	4	0.8
C2	The attacker can eavesdrop the communication between communicating stations, e.g., through sniffing the network.	4	0.8
C3	The attacker can remove, inject, or modify messages exchanged between communicating stations. This includes sending fabricated commands to the VC station.(C3 includes C2)	3	0.5
C4	The attacker can remotely access a VC station and change installed software, e.g., using an already installed malware.	2	0.2
C5	The attacker has physical access to control the use of the VC, e.g., power it on or off.	2	0.2
C6	The attacker can modify the hardware installed in the VC station and change its circuits (e.g., maintenance officer).	1	0.01

Table 7 – Selected set of threats to the system and associated attacker capabilities.

ID	Threat	Example of attack scenarios	Required capabilities
T1	Unauthorized join of meetings.	VC station has auto-answer feature on. The attacker sends connect request to the VC station and joins automatically the meetings—no authentication is required.	C1 or C4 or C6
T2	Unauthorized control of the movements of the camera.	The attacker guesses the authentication credentials (e.g., uses easy to guess passwords) of one of the users who has the authorization to control the movements of the camera.	C3 or C4 or C6
T3	Unauthorized access to frames exchanged between the VC station and communicating stations.	The attacker updates the VC station firmware to send him a copy of each frame sent to other VC stations.	C2 or C4
T4	Unauthorized modification of frames sent from a VC station to communicating station.	The attacker controls one of the routers used in the communication between both parties—e.g., an Internet service provider. He captures the frames exchanged between the VC stations, which are not signed, modifies them, and injects the modified frames to the receiving party.	C3 or C4 or C6
T5	Interrupt the VC station.	Assume that the VC does not implement a protection from Denial of Service (DoS) attacks (Gilgor, 1983). The attacker exploits the vulnerability and floods the VC station with high rate of messages until the station stops responding.	C3 or C4 or C5 or C6
T6	Unauthorized change of the behavior of the VC station.	Assume that updating the firmware of the VC station does not need authentication. The attacker exploits the weakness and uploads a modified firmware to the VC station, which changes the behavior of the device.	C4 or C6

and show how changing attacker capabilities could be used to mitigate threats.

6.1.2. *Estimating the security risks of the system*

Assume we identified the initial attacker capabilities and threats to the VC system using the risk assessment process described in Section 4.1. We apply the risk estimation method as follows. First, we evaluate for each attacker capability whether it is certain, easy, possible, unlikely or very unlikely that potential attackers can have the given capability. This requires (1) identification of entities that have the capability and (2) estimation of the chance that one or a set of the entities become attackers. We report the authors' own perception about the likelihood of a set of attacker capabilities corresponding to the VC system, as an illustration of the activity, in Table 6.[13]

Second, we identify for each threat one or many attack scenarios and the capabilities required to cause it. Table 7 lists a selected set of threats to the VC system and their related attacker capabilities.

Third, we evaluate for each threat the factors that measure the ease of causing it by using one of the capabilities and assign a score to each factor. We estimate also the occurrence of each threat using information about the possible motivations to cause them, such as for fun, for fame and for financial gains. Table 8 reports—as an illustration of the activity—the authors' evaluation about the scores of the factors for each threat.

The scores provided in Table 8 show that for some threats the scores for the same factor vary based on the attacker capability being considered. For instance an attacker who intends to cause "Interruption of the VC station" threat (i.e., threat T5) and has physical access to the VC station (i.e., capability C5) does not need time to figure out how to push the power off button of the VC system, neither expertise, knowledge, equipment, or window of opportunity—we score 8 for all the factors. In contrast, an attacker who plans to cause the same threat but uses the capability inject messages (i.e., has capability C3) needs time, expertise, knowledge, and tools to figure out how to craft a message of a command to the VC station to power it off.

[13] The table shows attacker capabilities scores, which are derived from the truth values as discussed in Subsection 3.3.

Table 8 – Evaluation of the likelihood factors of the selected set of threats to the system.

ID	Threat	Capability	Elapsed time	Specialist expertise	Knowledge of the system	Window of opportunity	Required equipment	Occurrence	Total score
T1	Unauthorized join of meetings.	C1	4	6	4	8	8	1	24 (30)
		C4	4	3	4	8	8	1	5.4 (27)
		C6	4	3	4	8	8	1	0.27 (27)
T2	Unauthorized use of the VC station.	C3	4	3	4	8	8	1	13.5 (27)
		C4	4	3	2	8	2	1	3.8 (19)
		C6	4	3	2	8	1	1	0.18 (18)
T3	Unauthorized access to frames exchanged between a VC station and communicating station.	C2	6	6	8	8	1	1	23.2 (29)
		C4	6	6	8	8	1	1	5.8 (29)
T4	Unauthorized modification of frames sent from a VC station to communicating station.	C3	4	6	8	8	8	1	17 (34)
		C4	4	3	2	1	8	1	3.6 (18)
		C6	4	3	2	1	8	1	0.18 (18)
T5	Interrupt the VC station.	C3	4	3	2	1	2	0.5	3 (12)
		C4	4	3	2	1	4	0.5	1.4 (14)
		C5	8	8	8	8	8	0.5	4 (40)
		C6	4	3	2	1	4	0.5	0.07 (14)
T6	Unauthorized change of the behavior of the VC station.	C4	4	3	2	1	2	1	2.4 (12)
		C6	4	3	2	1	2	1	0.12 (12)

Note: Values between parentheses are the sum of the likelihood factor scores without considering the capability likelihoods.

Table 9 – Evaluation of the severity of the selected set of threats to the system.

ID	Threat	Safety	Privacy	Financial	Operational	Reputation damage	Total score
T1	Unauthorized join of meetings.	0	3	4	0	4	11
T2	Unauthorized use of the VC station, e.g., control the movements of the camera.	0	4	4	1	4	13
T3	Unauthorized access to frames exchanged between a VC station and communicating station.	0	3	4	0	4	11
T4	Unauthorized modification of frames sent from a VC station to communicating station.	0	3	4	0	4	11
T5	Interrupt the VC station.	0	0	1	4	1	6
T6	Unauthorized change of the behavior of the VC station.	0	0	2	4	4	10

Table 10 – Estimation of the risk exposure of the selected set of threats.

ID	Threat	Likelihood	Severity	Risk
T1	Unauthorized join of meetings.	24	11	264
T2	Unauthorized use of the VC station, e.g., control the movement of the camera.	13.5	13	175.5
T3	Unauthorized access to frames exchanged between a VC station and communicating station.	23.2	11	255.2
T4	Unauthorized modification of frames sent from a VC station to communicating station.	17	11	187
T5	Interrupt the VC station.	4	6	24
T6	Unauthorized change of the behavior of the VC station.	2.4	10	24

Fourth, we evaluate for each threat the severity factors and we assign scores that indicate the size of expected losses and damage. Table 9 reports about the authors own perception about the factors for each threat.

Fifth, we compute the risk exposure of each threat using the likelihood and severity data. Table 10 provides the likelihood, severity, and risk exposure scores of the selected threats for the VC system.

Note that there are two combinations of attacker capabilities: (1) alternative capabilities, which we represent using the operator "or" and (2) joined capabilities, which we represent using the operator "and."

6.1.3. *Risk reduction by changing attacker capabilities*

This section provides two examples that show how changing attacker capabilities reduces the risk of the system. Their description follows.

Example 1. Attackers wishing to cause threat *T1* (unauthorized join of meetings) may use capability *C1* (the attacker can discover the VC station). If we change the system architecture such that the VC stations do not reply to probing messages, then attacker capability *C1* becomes obsolete. Therefore, the attackers can cause the threat using capabilities *C4* or *C6*, but not *C1*. The change reduces the risk exposure of threat *T1* from score 264 (Table 10) to score 59.4.

Example 2. Attackers wishing to cause threat *T5* (interrupt the VC station) may use capability *C5* (physical access to control the use of the VC system). If we remove the buttons of the VC station and give the responsibility of using the remote control to the room manager, we make capability *C5* obsolete. Now, the attackers can still cause the threat using capabilities *C3*, *C4* or *C6*, but not *C5*. This change reduces the risk exposure of threat *T5* from score 24 (see Table 10) to score 18.

6.2. *Example 2: connected vehicles*

This section illustrates the use of our proposed approach using the example system-Connected Vehicles. First, we describe the system and then we estimate the risk exposures associated to the system.

6.2.1. *Description of the system*

A connected vehicle, as the one shown in Fig. 4, has a set of sensors and ECU) that use an in-vehicle network to communicate and control various operations of the vehicle. Modern connected vehicles have the ability to connect with personal devices, road side units and communicate with neighboring vehicles, service centers (ben Othmane et al., 2013b). This is used in several applications such as eCall, fleet management, and remote firmware update.

Besides traditional physical attacks, connected vehicles, because of their features, provide attackers the option to remotely connect with the vehicle and conduct attacks (ben Othmane et al., 2013b). Connected vehicles are prone to cyber-threats, for instance, an attacker could connect to the in-vehicle network of a connected vehicle, remotely update an ECU to disrupt its braking system, and make it crash with other vehicles. In the following we illustrate the use of the proposed method for estimating the risk exposures of threats to connected vehicles and show how changing attacker capabilities could be used to mitigate threats.

6.2.2. *Estimating the security risk exposures of the system*

We studied typical connected vehicle architectures to identify a set of potential threats. Then, we selected a subset of these threats which we believe are highly significant in the context of modern vehicles (Ruddle, March 2010). Based on these threats, we enumerated a set of possible attacker capabilities that are essential for an attacker to realize these threats. Table

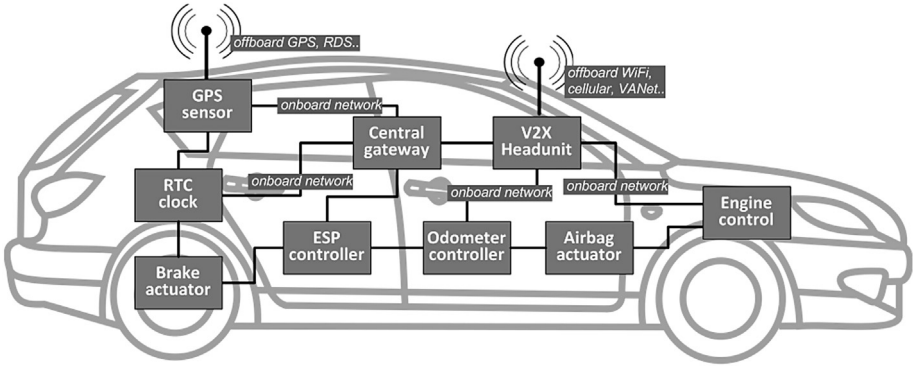

Fig. 4 – Architecture of connected vehicle (ben Othmane et al., 2015).

Table 11 – Sample of attacker capabilities corresponding to the system.

ID	Capability	Score	Likelihood
C1	Attacker can physically access the CAN bus (e.g. connect a new ECU to the CAN bus)	3	0.5
C2	Attacker can remotely inject messages to CAN bus	2	0.2
C3	Attacker can spoof external GPS signals	4	0.8
C4	Attacker can control communication between the vehicle and the Internet	2	0.2

Table 12 – Selected set of threats to the system and associated attacker capabilities.

ID	Threat	Example of attack scenarios	Required capabilities
T1	Falsification of speedometer reading of the vehicle	An attacker may alter the speedometer reading seen by the driver, which may cause the driver to make wrong driving decisions.	C1 or C2
T2	Disruption of the braking system of the vehicle	An attacker may disable the breaking system while the car is in motion, or apply breaks when the driver doesn't expect it.	C1 or C2
T3	Disruption of the emergency response system of the vehicle (e.g., OnStar)	Some modern vehicles are equipped with emergency response systems, where the driver and passengers can contact some party to request assistance in emergency situations. An attacker may completely disable this system or falsify any information provided by the system.	C1 or C3 or C4
T4	Generating false check lights in the dashboard on the vehicle	Drivers depend on information displayed in the dashboard for warnings such as low tire pressure and low fuel level. An attacker may alter this information to trick the driver into driving the car until it runs out of fuel or making him/her pull over due to a false tire pressure warning.	C1 or C2
T5	Locking the gearstick in a fixed position	An attacker can use such an attack to render the vehicle immobile.	C1 or C2
T6	Sending deceptive messages to the infotainment system	An attacker sends wrong GPS information on behalf of a neighboring vehicle, so it receives non desired infotainment service.	C1 or C3 or C4
T7	Remotely update an ECU	Attacker may update an ECU of the vehicle with malicious firmware forcing the vehicle to misbehave.	C2

11 lists the identified capabilities. Table 12 describes the selected threats and shows the related capabilities.

We conducted a questionnaire to obtain expert opinions on the set of capabilities and threats we identified, which we describe in details in Section 7. First, the experts rated each capability in terms of the likelihood that an attacker may obtain that capability. The possible options were "Impossible", "Very Unlikely", "Likely", "Possible", "Highly Likely", "Certain/Sure." Next, the experts rated the factors listed in Table 4 for each threat. Finally, they rated these factors with respect to each applicable capability of each threat.

Table 13 reports about the authors own evaluation of the likelihood factors of the selected set of threats to the system, Table 14 reports about the authors own perception about the severity of each threat. Table 15 provides the risk exposure score of each of the selected threats.

6.2.3. *Risk exposure reduction by changing attacker capabilities*

This section provides two examples that show how changing attacker capabilities reduces the risk of the system. Their description follows.

Example 1. Attackers wishing to cause threat *T3* (disruption of the emergency response system of the vehicle) may use capability *C3* (the attacker can spoof external GPS signals). If we change the system architecture such that the connected vehicles use a positioning system resistant to spoofing (Jafarnia-Jahromi et al., 2012), capability *C3* becomes obsolete. Therefore, the attackers can cause the threat using capabilities *C1* or *C4*, but not *C3*. The change reduces the risk exposure of threat *T3* from score 168 (Table 15) to 91.

Example 2. Attackers wishing to cause threat *T6* (sending deceptive messages to the infotainment system) may use capability *C3* (the attacker can spoof external GPS signals). If we change the system architecture such that the connected vehicles use a positioning system resistant to spoofing (Jafarnia-Jahromi et al., 2012), capability *C3* becomes obsolete. Therefore, the attackers can cause the threat using capabilities *C1* or *C4*, but not *C3*. The change reduces the risk exposure of threat T6 from score 144 (Table 15) to 84.

7. Evaluation of the proposed risk estimation method

This section evaluates whether considering attacker capabilities helps the security experts to provide more accurate risk estimates or not. In the following we describe the evaluation method that we use (Subsection 7.1), describe and report the results of two experiments we use to evaluate the method (Subsection 7.2), and conclude the evaluation (Subsection 7.3).

7.1. *Uncertainty evaluation method*

There are two approaches for evaluating the estimation errors: the error approach and the uncertainty approach. The error approach measures the difference between the estimated values and the corresponding presumed true values.

Table 13 – Evaluation of the likelihood factors of the selected set of threats to the system.

ID	Threat	Capability	Elapsed time	Specialist expertise	Knowledge of the system	Window of opportunity	Required equipment	Occurrence	Total score
T1	Falsification of speedometer reading of the vehicle	C1	6	6	4	2	8	1	13 (26)
		C2	4	6	4	2	8	1	4.8 (24)
T2	Disruption of the braking system	C1	6	6	4	2	8	1	13 (26)
		C2	4	1	4	2	8	1	3.8 (19)
T3	Disruption of the emergency response system of the vehicle (e.g., OnStar)	C1	6	6	4	2	8	1	13 (26)
		C3	8	8	4	2	8	1	24 (30)
		C4	8	8	2	2	8	1	5.6 (28)
T4	Generating false check lights in the dashboard on the vehicle	C1	6	6	4	2	8	1	13 (26)
		C2	4	1	4	2	8	1	3.8 (19)
T5	Locking the gearstick in a fixed position	C1	6	6	4	2	8	1	13 (26)
		C2	4	6	4	2	8	1	4.8 (24)
T6	Sending deceptive messages to the infotainment system	C1	6	8	4	2	8	1	14 (28)
		C3	8	8	4	2	8	1	24 (30)
		C4	8	8	4	2	0	1	4.4 (22)
T7	Remotely update an ECU	C2	2	1	2	2	0	1	1.4 (7)

Note: Values between parentheses are the sum of the likelihood factor scores without considering the capability likelihoods.

Table 14 – Evaluation of the severity of the selected set of threats to the system.

ID	Threat	Safety	Privacy	Financial	Operational	Reputation damage	Total score
T1	Falsification of speedometer reading of the vehicle	3	0	0	1	3	7
T2	Disruption of the braking system of the vehicle	4	0	0	2	4	10
T3	Disruption of the emergency response system of the vehicle (e.g., OnStar)	0	0	2	2	3	7
T4	Generating false check lights in the dashboard on the vehicle	0	0	0	1	0	1
T5	Locking the gearstick in a fixed position	4	0	0	2	3	9
T6	Sending deceptive messages to the infotainment system	0	0	2	2	2	6
T7	Remotely update an ECU	4	0	0	4	4	12

Table 15 – Estimation of the risk exposure of the selected set of threats.

ID	Threat	Likelihood	Severity	Risk exposure
T1	Falsification of speedometer reading of the vehicle	13	7	51
T2	Disruption of the braking system of the vehicle	13	10	130
T3	Disruption of the emergency response system of the vehicle (e.g., OnStar)	24	7	168
T4	Generating false check lights in the dashboard on the vehicle	13	1	13
T5	Locking the gearstick in a fixed position	13	9	121
T6	Sending deceptive messages to the infotainment system	24	6	144
T7	Remotely update an ECU	1.4	12	16.8

The uncertainty approach measures the bounds within which a true value may be reasonably presumed to lay (Birch, Mar. 2003).

The use of the error approach in estimating the risk exposure of threats is not feasible because of the difficulty to get historical data about threat occurrence frequencies—which could be considered the true values. We adopt the second approach—which is feasible—to evaluate the method we propose for estimating the ease of causing threat (defined in Section 4.2).

The uncertainty approach uses the standard deviation as a metric of uncertainty (Working Group 1 of the Joint Committee for Guides in Metrology (JCGM/WG 1), 2012) (Birch, Mar. 2003) (Physical Measurement Laboratory of NIST, Oct. 2000).[14] The approach is commonly used in Physics and Chemistry to measure the uncertainty of the results of the experiments. Measuring the uncertainty of the results allows accounting for the conditions of the given experiment, such as the precision of the measurement instruments.

Note that we use the term *uncertainty*, which is commonly used when evaluating decisions with incomplete information and lack of knowledge (Agarwal et al., 2004), such as in Dempster-Shafer evidence theory (Shafer, 1976) and fuzzy logic (Zadeh, 1975). We use the term uncertainty in this paper to mean a measurement that "characterizes" the dispersion of the values of a sample (Working Group 1 of the Joint Committee for Guides in Metrology (JCGM/WG 1), 2012).

This uncertainty metric, in the context of our experiments, measures the degree to which the elements within the sample differ from the sample mean. The metric is based on the assumption that if a set of experts provide close estimates for the ease of causing a given threat, then the estimates are considered accurate, otherwise it indicates how much we are uncertain about the accuracy of the estimates.[15]

$$z_1 = \sum_{i=1..5} x_i \tag{9}$$

$$z_2 = y \times \sum_{i=1..5} x_i \tag{10}$$

$$\begin{aligned} \Delta z_1 &= \sqrt{\sum_{i=1..5} \left(\frac{\partial z_1}{\partial x_i} \times \Delta x_i \right)^2} \\ &= \sqrt{\sum_{i=1..5} (\Delta x_i)^2} \end{aligned} \tag{11}$$

The score of ease of causing threat that we aim to measure combines the scores of a set of factors—see section 4.2. Equation (9) formulates the ease of causing threat without considering attacker capabilities in terms of the 5 factors of Table 4, represented by variables x_i. Equation (10) formulates the ease of causing threat considering attacker capabilities in terms of the 5 factors represented by variables x_i and capability y.[16]

Equation (11) formulates computing the uncertainty of ease of causing threat without considering attacker capabilities—computed using z_1—and Equation (12) formulates computing the uncertainty of ease of causing threat considering attacker capabilities—computed using z_2. The equations use Andraos' derivation formula for computing uncertainty using partial derivatives of the variables used to compute z_1 and z_2 (Andraos, 1996).[17]

$$\begin{aligned} \Delta z_2 &= \sqrt{\sum_{i=1..5} \left(\frac{\partial z_2}{\partial x_i} \times \Delta x_i \right)^2 + \left(\frac{\partial z_2}{\partial y} \Delta y \right)^2} \\ &= \sqrt{\sum_{i=1..5} (\bar{y} \times \Delta x_i)^2 + \left(\left(\sum_{i=1..5} \bar{x}_i \right) \times \Delta y \right)^2} \end{aligned} \tag{12}$$

7.2. Evaluation experiments

The goal of the study is to investigate whether using attacker capabilities improves the accuracy of the expert estimates of likelihood of the ease of causing threats or not. The study addresses the following research question: Does incorporating attacker capabilities reduce the uncertainty in the experts opinions about the likelihood of the ease of causing threats?

We describe and report in the following the results about two quantitative analysis experiments (Wohlin et al., 2012) that we conducted to answer this research question. (Both experiments address the same research question.)

[14] Note that the NIST defines uncertainty for a single measurement (Physical Measurement Laboratory of NIST, Oct. 2000) while Birch defines uncertainty for a sample of measurements (Birch, Mar. 2003). We use the latter in this work.

[15] This is based on the common social behavior to believe in facts that the experts agree upon and have uncertainty about facts that the experts provide different opinions about them.

[16] The equation is equivalent to Equation (2).

[17] We have a class Type B of uncertainty evaluation (Working Group 1 of the Joint Committee for Guides in Metrology (JCGM/WG 1), 2012) (Birch, Mar. 2003).

7.2.1. Experiments planning

This section describes the planning of evaluation experiments. The planning includes the hypothesis to be tested by the experiments, dependents and independent variables, experiments design, and instrumentation.

Hypothesis. The objective of the study is to check whether considering attacker capabilities in risk estimation reduces the uncertainty of the likelihood of the ease of causing threats or not. We test the following hypothesis:

> The uncertainty in the ease of causing threat when not considering attacker capability is equal to the uncertainty of ease of causing threat when considering attacker capabilities.

7.2.1.1. Variables. The variables that we want to study are called dependent variables and the variables that we can control/manipulate are called independent variables (Wohlin et al., 2012). In this experiment, the independent variables are the likelihood estimation factors provided in Subsection 4.2. They are: elapsed time, specialist expertise, knowledge of the system, window of opportunity, and required equipment and tools. The dependent variables are: (1) the uncertainty of ease of causing threat without considering attacker capabilities and (2) the uncertainty of ease of causing threat considering attacker capabilities. The formulas to compute both dependents variables are provided in Subsection 7.1.

7.2.1.2. Design. We developed 2 questionnaires to test whether incorporating attacker capabilities reduces the uncertainty that experts have in evaluating the likelihood of the ease of causing threats or not. We used the questionnaires to collect expert opinions about the likelihood of ease of causing threats for hypothetical IS: a VC system and connected vehicles. Each questionnaire includes two parts:

1. Part 1—Estimating the likelihood of ease of causing threats without considering attackers capabilities. In this part each security expert evaluates for each threat the factors that measure the ease of causing the threat.
2. Part 2—Estimating the likelihood of ease of causing threats considering attackers capabilities. In this part each expert estimates the likelihood of each attacker capability and evaluates for each threat the factors for measuring the ease of causing the threat considering the applicable attacker capabilities.

We used Equations (11) and (12) to compute the uncertainty in the measurement ease of causing threat for each threat for both questionnaire parts. Then, we tested the hypothesis of the study/experiments.

We discuss in the following the conduct of the experiments.

7.2.2. Experiments operation

The selection criteria of participants for both experiments are: (1) experience with security attacks, (2) experience with risk assessment, and (3) good knowledge about the domain of the system object of the study. We discuss in the following the conduct of the 2 experiments.

7.2.2.1. Experiment 1: video conferencing system. We sent requests to a set of domain experts and security experts in October 2013. The domain experts were software developers of a video conferencing system, which are trained to develop secure software—including doing threat modeling and risk estimation. The security experts where security researchers who are familiar with the system. The questionnaire was available for the participants for few weeks.[18] We accepted the data of 6 full participants: 3 security experts and 3 domain experts[19]

7.2.2.2. Experiment 2: connected vehicles. We sent invitations to a set of security experts to participate in the study in November 2013. The experts were security researchers who are familiar with connected vehicles. The questionnaire was available for the participants for few weeks.[20] We received 9 full participations in this study.

7.2.3. Data analysis and interpretation

This subsection reports about the analysis we performed on the data collected using both experiments.

7.2.3.1. Experiment 1: video conferencing system. We computed the mean and the uncertainty of the expert estimates for each threat for both parts of the questionnaire. Table 16 reports the estimates of the capability likelihoods and Table 17 reports the means and uncertainty of the threat causing ease likelihoods considering and not considering attacker capabilities.

Table 17 shows that the mean of the uncertainty of the likelihoods of the ease of causing threats without considering attacker capabilities is 4.60 and the mean of the uncertainty of likelihoods of the ease of causing threats considering attacker capabilities is 2.29. The t-test concludes that there is a significant difference in the scores for considering attacker capabilities ($M = 2.29$, $SD = 0.28$) and not considering attacker capabilities ($M = 4.60$, $SD = 0.26$) conditions; $t(5) = 14.41$, p-value $= 2.898e\text{-}05$. Moreover, the effect test concludes that the difference is of practical significance; the value of the effect size is 8, which is more than the threshold 0.33. These results suggest that incorporating attacker capabilities in estimating the likelihoods of the ease of causing threats reduces the uncertainty of the estimates.

Table 17 indicates that threat *T1* (Unauthorized join of meetings) has the highest priority and threat *T6* (Unauthorized change of the behavior of the VC station) and *T2* (Unauthorized use of the VC station) have the lowest priority if we do not consider attacker capabilities in estimating the ease of causing threats, and threat *T3* (Unauthorized access to frames

[18] This data collection was carried out with the Purdue University IRB authorization (Exemption #1309014017).

[19] We received input from 10 participants. We discarded the data of 4 participants because e.g., the data is incomplete.

[20] This data collection was carried out with the Purdue University IRB authorization (Exemption #1310014174).

Table 16 – Estimate of capability likelihoods. (Average of the experts' estimates).

ID	Attacker capability	Mean	STD
C1	The attacker can discover the VC station and the station has automatic reply feature on.	0.42	0.23
C2	The attacker can eavesdrop the communication between communicating stations	0.56	0.17
C3	The attacker can remove, inject, or modify messages exchanged between communicating stations	0.33	0.15
C4	The attacker can remotely access a VC station and change installed software	0.31	0.16
C5	The attacker has physical access to control the use of the VC	0.36	0.29
C6	The attacker can modify the hardware installed in the VC station and change its circuits	0.31	029

exchanged between a VC station and communicating station) has the highest priority and threat *T2* (Unauthorized use of the VC station) has the lowest priority if we consider attacker capabilities.

We believe that the threat ranking when considering attacker capabilities is more realistic. For instance, we believe that Threat *T3* (Unauthorized access to frames exchanged between a VC station and communicating station) is the easiest threat to cause because it is easy for potential attackers to eavesdrop the communication link between the communicating stations (capability *C2*) and use it to cause the threat. Also, it is difficult to cause threat *T1* (Unauthorized join of meetings) because that requires the ability to change the behavior of the target VC station (capabilities *C4* and *C6*) or to discover the VC station assuming it has the automatic reply feature on (capability *C1*).

7.2.3.2. Experiment 2: connected vehicles. We computed the mean and the uncertainty of the expert estimates for each threat for both parts of the questionnaire. Table 18 reports the estimates of the capability likelihoods and Table 19 reports the means and uncertainty of the threat causing ease likelihoods considering and not considering attacker capabilities.

Table 19 shows that the mean of the uncertainty of the likelihoods of the ease of causing threats without considering attacker capabilities is 5.20 and the mean of the uncertainty of

Table 18 – Estimate of capability likelihoods. (STD stands for Standard Deviation).

ID	Attacker capability	Mean	STD
C1	Attacker can physically access the CAN bus (e.g. Connect a new ECU to the CAN bus)	0.46	0.23
C2	Attacker can remotely inject messages to CAN bus	0.46	0.21
C3	Attacker can spoof external GPS signals	0.62	0.22
C4	Attacker can control communication between the vehicle and the Internet.	0.62	0.26

likelihoods of the ease of causing threats considering attacker capabilities is 2.23. The t-test concludes that there is a significant difference in the scores for considering attacker capabilities ($M = 2.23$, $SD = 0.09$) and not considering attacker capabilities ($M = 5.20$, $SD = 0.49$) conditions; $t(5) = 14.17$, p-value = 3.143e-05. Moreover, the effect size test concludes that the difference is of practical significance; the value of the effect size is 2 which is more than the threshold 0.33. These results suggest that incorporating attacker capabilities in estimating the likelihoods of the ease of causing threats reduces the uncertainty of the estimates.

Table 19 indicates that if we do not consider attacker capabilities the 5 threats have close ease of causing threat likelihood and threat *T4* (Generating false check lights in the dashboard on the vehicle) has the highest priority. The table indicates also that if we consider attacker capabilities there are two groups of threats based on the closeness of their average ease of causing threat likelihood. Threat *T3* (Disruption of the emergency response system of the vehicle (e.g., OnStar)) and threat *T6* (Sending deceptive messages to the infotainment system) form the group with highest priority, group *A*, and the remaining threats form the group with the lowest priority, group *B*.

We believe that the threat ranking when considering attacker capabilities is more realistic; that is, it is easier to cause the threats *T3* and *T6* than the others. The reason is that it is easier for attackers to communicate remotely with a target vehicle than to have physical access to its internal network—e.g., its Controller Area Network (CAN) bus and the tools and knowledge required to cause the threats are becoming available—see e,g., (Miller and Valasek, 2014). For instance, Checkoway et al. (Checkoway et al., 2011)

Table 17 – Estimate of likelihoods of ease of causing threats considering and not considering attacker capabilities.

ID	Threats	Not considering attacker capabilities		Considering attacker capabilities	
		Mean	Uncertainty	Mean	Uncertainty
T1	Unauthorized join of meetings	20.67	4.58	8.68	2.20
T2	Unauthorized use of the VC station	14.83	4.68	5.97	2.02
T3	Unauthorized access to frames exchanged between a VC station and communicating station	19.00	4.97	10.50	2.18
T4	Unauthorized modification of frames sent from a VC station to communicating station	15.33	4.56	6.68	2.21
T5	Interrupt the VC station	20.33	4.69	10.17	2.85
T6	Unauthorized change of the behavior of the VC station	14.83	4.15	6.62	2.28
Mean		4.60		2.29	

Table 19 – Estimate of likelihoods of ease of causing threats considering and not considering attacker capabilities.

ID	Threats	Not considering attacker capabilities		Considering attacker capabilities	
		Mean	Uncertainty	Mean	Uncertainty
T1	Falsification of speedometer reading of the vehicle	19.82	4.62	9.43	2.30
T2	Disruption of the braking system of the vehicle	18.27	5.53	8.43	2.17
T3	Disruption of the emergency response system of the vehicle (e.g., OnStar)	18.55	5.68	11.57	2.36
T4	Generating false check lights in the dashboard on the vehicle	20.82	4.88	7.97	2.11
T5	Locking the gearstick in a fixed position	18.73	5.70	8.51	2.15
T6	Sending deceptive messages to the infotainment system	18.27	4.76	11.37	2.28
Mean		5.20		2.23	

demonstrated a set of remote attacks on a connected vehicle that has e-call application.

7.2.4. *Validity of the study*

This section discusses the limitations of the validity of the study and the measures we took to control them. The classes of validity are: conclusion validity, internal validity, construct validity, and external validity (Wohlin et al., 2012).

7.2.4.1. *Conclusion validity*. This validity concerns the relationship between each experiment and the results of the related data analysis. We addressed this validity using three measures. First, we tested the hypothesis using two questionnaires, each evaluates the hypothesis for a specific IS example. Second, since the sizes of the samples of both tests are limited, we used the student distribution to infer our results. The student distribution is used in situations where the sample size, drawn from a normal distribution, is small (Gosset, 1908). We also used the effect size to test whether the difference between two sample means (in this case the mean of the likelihood uncertainties) is of practical consequence. A low effect size indicates that we shall not derive conclusions from the results and a high effect size indicates that the results could be used to derive conclusions. The effect size value is compared to the threshold 0.33, which is a consensus value.[21] Third, we targeted participants who are supposed to be security experts. Unfortunately, since the questionnaire was online and data collection was anonymous, we may had participants who were not security experts.[22]

7.2.4.2. *Internal validity*. This validity concerns the causal relationship between the experiments and the results of the analysis. There are three limitations to the internal validity of the experiments. First, the questionnaire for Example 1 invites the participants to watch a video that shows attacks that apply to the system, which may impact the opinions of the experts. Second, the experiment results could be affected by the fact that each participant must take successively the two parts of the questionnaire, and the hypothesis compares the data of these parts. Third, the experiment results could be affected by the quality of the questions. We addressed this threat by testing the questionnaire before making them available online.

7.2.4.3. *Construct validity*. This validity concerns the relation between the experiments and the hypothesis and between the experiments and the results of the analysis. There are 2 limitations to the validity of the study. The first is the difference between perception and reality in questionnaires (Likert). The second is the choice of independent variables; we used, in the experiments, a set of factors for estimating the likelihood of threats that are commonly used, but their effectiveness is not assured (We discuss the issue in the next section.).

7.2.4.4. *External validity*. This validity concerns the condition to the generalization of the results. The study could be generalized further; we already tested the hypothesis using two experiments, each evaluates the hypothesis for a specific IS example.

7.3. *Summary and conclusions*

This section focuses on evaluating whether incorporating attacker capabilities reduces the uncertainty in the experts opinions about the likelihood of the ease of causing threats or not. The evaluation is performed using a quantitative analysis experiments that tests the hypothesis: The uncertainty in the ease of causing threat when not considering attacker capability is equal to the uncertainty of ease of causing threat when considering attacker capabilities.

The experiments setting includes two online questionnaires that are used to collect expert opinions about the likelihood of ease of causing threats for hypothetical IS: a VC system and connected vehicles. We tested the hypothesis of both systems. The results suggest that considering attacker capabilities enables experts to be more certain in estimating the likelihoods of ease of causing the threats to the systems they analyze.

We believe that the obtained results are sufficient evidence to support our claim. The main limitation to the validity of the experiments is the small sample sizes. The limitation is addressed through the use of statistical metrics based on the student distribution (Gosset, 1908) and effect test. The use of the student distribution enables hypothesis testing using small size samples and the effect test allows evaluating whether the results are of practical significance or not. We also used two examples to check the hypothesis to strengthen the validity of our conclusions.

[21] See for example http://www.fgse.nova.edu/edl/secure/stats/lesson5.htm.

[22] We used authentication code for the questionnaire, which should help to control this threat.

8. Impacts of the results on the state of the art

The paper proposes incorporating attacker capabilities in the risk exposure estimates of threats to ISs, which extends the ISSRM domain model proposed by Dubois et al. (Dubois et al., 2010) and existing approaches for risk estimation, such as the model proposed by Wheeler (Wheeler, 2011). The paper argues that (1) denying attacker capabilities could be used to mitigate the risk to ISs and (2) considering attacker capabilities in estimating the likelihoods of ease of causing the threats to ISs helps reducing the uncertainty of the estimates.

Attacker capabilities are special conditions for threats; making the conditions impossible implies mitigating the risks of the related threats. Therefore, one approach to mitigate the risk of a given threat is denying attackers the required capabilities for the threat, when possible—the approach is not feasible for all threats to ISs. This could be achieved through changing the architecture or the access policies of the systems. Examples of mitigating risks of threats through denying attacker capabilities are provided in Section 6.

This result constitutes a new approach for risk mitigation. Managers of ISs could consider mitigating threats by changing the architecture and the access policies of their systems as an alternative to implementing security mechanisms—when possible. We believe that for some cases such change may provide better return on investment.

The paper also demonstrates through 2 quantitative experiments that considering attacker capabilities in estimating the likelihoods of ease of causing the threats to ISs helps reducing the uncertainty of the estimates—see Section 7. The results suggest providing the experts with the extra information, attacker capabilities, helps them to be more certain in estimating the likelihoods of ease of causing the threats to the system they analyze. Unfortunately, incorporating attacker capabilities in estimating the risk exposures of threats creates an overhead. However, this limitation could be controlled by limiting, for example, the attacker capabilities to be considered for each threat to the 3 capabilities that have the highest likelihoods.

Recall that the high uncertainty in the risk estimates leads the business managers to view the risk estimation activity as of uncertain practical value (Bonnette, July 2003; Wheeler, 2011). The finding (incorporating attacker capabilities in estimating the likelihoods of ease of causing the threats to ISs helps reducing the uncertainty of the estimates) helps producing risk estimation with less uncertainty. This shall reduce the negative perception of the business managers about the practical value of the threat estimation activity.

Moreover, current methods of security risk estimation rely on expert judgments to estimate the risk exposures of threats using estimation factors that vary based on the used estimation method. The effectiveness of these proposed risk estimation models have not been objectively evaluated (cf. (Hubbard, 2009)). The method that we proposed and used in this paper to measure the uncertainty of security experts in estimating the likelihood of ease of causing threat could be used in quantitative analysis experiments (Wohlin et al., 2012) to evaluate the effectiveness of the threat likelihood factors and risk estimation methods.

Finally, the paper shows that incorporating attacker capabilities reduces the uncertainty in estimating the ease of causing threats and therefore the risk exposure estimates. This leads to the questions: What are the information that experts need to use so that the security risk exposure estimates can be improved? How to measure the added value of the information used in the security risk exposure estimates? And what is the threshold of the number of factors we should consider in estimating the risk exposures of threats? We will investigate these questions in our future work.

9. Conclusions

The goal of the risk estimation activity is to enable business owners of ISs to prioritize addressing the threats. The high uncertainty in the risk exposure estimates, indicated by the high differences between the estimates, leads the business managers to view the exercise as of uncertain practical value.

This paper proposes incorporating attacker capability in estimating the likelihoods of ease of causing the threats to ISs, which is one of the risk estimate components. An attacker capability is the ability to access or use a set of resources of the ISs to exercise threats to the system. It enables the use of appropriate means and opportunities to cause a given threat. The paper proposes a method for incorporating attacker capabilities in estimating the likelihood of threats, illustrates the method through two examples: a video conferencing system and connected vehicles system, demonstrates that denying attacker capabilities can reduce security risk to IS, and evaluates the effect of considering attacker capabilities on the uncertainty in estimating the likelihoods of threats. It also shows empirically that considering attacker capabilities in estimating the likelihoods of ease of causing the threats to ISs helps reducing the uncertainty of the estimates.

The use of attacker capabilities in estimating the likelihood of ease of causing threats impacts the advices about the priority of mitigating threats. We suggest to consider attacker capabilities in estimating the likelihood of ease of causing threats because it helps to reduce the uncertainty of estimating the risk exposures of threats to ISs, be more confident in ranking the threats, and have better return on investment in developing security measures for software.

The results of this work suggests the following 2 future work directions. First, the identification of the impacts of attack capabilities on the likelihood of ease of causing threats and risk explore estimate of threats in general leads to several other questions, such as what are the information that experts need to use so that the uncertainty in security risk exposure estimates can be reduced? Second, the inherent uncertainty of threat likelihood estimates due to the fact that experts provide estimates based on perception rather than reality leads to the question: How to account for such uncertainty in estimating risks of IS threats?

Acknowledgments

This work was supported, in part, by the BMBF within EC SPRIDE (01BY1171), the Hessian LOEWE excellence initiative within CASED, and a Fraunhofer Attract (project 600130) grant. The authors thank the anonymous participants in the study and Andreas Poller, Philipp Holzinger, Stefan Triller, and Jan Steffan from Fraunhofer SIT for their valuable comments and discussions.

REFERENCES

Agarwal H, Renaud J, Preston E, Padmanabhan D. Uncertainty quantification using evidence theory in multidisciplinary design optimization. Reliab Eng Syst Saf 2004;85(13):281–94. Alternative Representations of Epistemic Uncertainty.

Alberts CJ, Dorofee A. Managing information security risks: the Octave approach. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.; 2002.

Andraos J. On the propagation of statistical errors for a function of several variables. J Chem Educ 1996;73(2):150. http://dx.doi.org/10.1021/ed073p150.

ANSSI. Expression des besoins et dentification des objectifs de sécurité (ebios)—bases de connaissances. Jan. 2010. URL, http://www.ssi.gouv.fr/IMG/pdf/EBIOS-2-BasesDeConnaissances-2010-01-25.pdf.

Apostolakis GE. How useful is quantitative risk assessment? Risk Anal 2004;24(3):515–20. http://dx.doi.org/10.1111/j.0272-4332.2004.00455.x. URL, http://dx.doi.org/10.1111/j.0272-4332.2004.00455.x.

ben Othmane L, Weffers H, Klabbers M. Using attacker capabilities and motivations in estimating security risk. In: Workshop on risk perception in it security and privacy, Newcastle, UK; 2013. URL, http://cups.cs.cmu.edu/soups/2013/risk/Cap.-Based-risk.pdf.

ben Othmane L, Al-Fuqaha A, ben Hamida E, van den Brand M. Towards extended safety in connected vehicles. In: Proc. 16th international IEEE conference on intelligent transportation systems, the Hague, The Netherlands; 2013. p. 652–7.

ben Othmane L, Weffers H, Mohamad MM, Wolf M. Wireless sensor networks (WSN) for vehicular and space applications: architecture and implementation. Norwell, MA: Springer; 2015. Ch. A Survey of Security and Privacy in Connected Vehicles, inPress.

Birch K. Estimating uncertainties in testing: an intermediate guide to estimating and reporting uncertainty of measurements in testing, british measurement and testing association. Meas Good Pract Guide Mar. 2003;36. URL, http://www.dit.ie/media/physics/documents/GPG36.pdf.

Bishop M. Computer security: art and science, Addison-Wesley. 2012. 13th printing.

Bishop M, Gates C. Defining the insider threat. In: Proc. Of the 2008 cyber security and information infrastructure research workshop, Oak Ridge, TN; 2008.

Bishop M, Engle S, Frincke D, Gates C, Greitzer F, Peisert S, et al. A risk management approach to the "insider threat". In: Probst CW, Hunker J, Gollmann D, Bishop M, editors. Insider threats in cyber security. Vol. 49 of Advances in Information Security, Springer US; 2010. p. 115–37.

Boehm B. Software risk management: principles and practices. Softw IEEE 1991;8(1):32–41. http://dx.doi.org/10.1109/52.62930.

Bonnette CA. Assessing threats to information security in financial institutions. July 2003. URL, https://www.sans.org/reading-room/whitepapers/threats/assessing-threats-information-security-financial-institutions-1143.

BSI. Threats catalogue elementary threats. August 2012. URL, https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Grundschutz/download/threats_catalogue.pdf?__blob=publicationFile.

Checkoway S, McCoy D, Kantor B, Anderson D, Shacham H, Savage S, et al. Comprehensive experimental analyses of automotive attack surfaces. In: Proc. of the 20th USENIX conference on Security, Berkeley, CA; 2011. 6–6.

Chivers H, Clark JA, Cheng P-C. Risk profiles and distributed risk assessment. Comput Secur 2009;28(7):521–35.

Clements P, Bachmann F, Bass L, Garlan D, Ivers J, Little R, et al. Documenting software Architectures: Views and beyond. 2nd ed. Addison Wesley; 2011.

Dictionary M. Macmillan dictionary. accessed on Feb. 2014. 2014. URL, http://www.macmillandictionary.com/us/dictionary/american/capability.

Dubois E, Heymans P, Mayer N, Matulevicius R. A systematic approach to define the domain of information system security risk management. In: Nurcan S, Salinesi C, Souveyet C, Ralyte J, editors. Intentional perspectives on information systems engineering, Springer Berlin Heidelberg; 2010. p. 289–306.

Duggan DP, Thomas SR, Veitch CKK, Woodard L. Categorizing threat—building and using a generic threat matrix. Tech. Rep. SAND2007–5791. USA: Sandia National Laboratories; Sep. 2007.

Ekelhart A, Fenz S, Neubauer T. Aurum: a framework for information security risk management. In: System sciences, 2009. HICSS '09. 42nd Hawaii International conference on; 2009. p. 1–10. http://dx.doi.org/10.1109/HICSS.2009.82.

Gilgor V. A note on the denial-of-service problem. In: Proc. Of the IEEE symposium on security and privacy, SP '83, Oakland, CA; 1983. p. 139–49.

Gosset W. The probable error of a mean. Biometrika 1908;6(1):1–25 (Student).

Hernan S, Lambert S, Ostwald T, Shostack A. Uncover security design flaws using the stride approach. MSDN Mag November 2006. URL, http://msdn.microsoft.com/en-us/magazine/cc163519.aspx.

Howard M, Lipner S. The security development Lifec. Redmond, WA: Microsoft Press; 2006. p. 101–32. Ch. Stage 4: Risk Analysis.

Hu VC, Ferraiolo DF, Kuhn DR. Assessment of access control systems. Tech. Rep. NISTIR 7316. Gaithersburg, MD: National Institute of Standards and Technology (NIST); Sep. 2006.

Hubbard D. The failure of risk management: Why it's broken and how to fix it. Hoboken, NJ: John Wiley and Sons; 2009.

ISO/IEC 18045. Information technology security techniques methodology for it security evaluation. 2008.

ISO/IEC 27002. Information technology security techniques code of practice for information security management. 2005. URL, http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=50297.

Jafarnia-Jahromi A, Broumandan A, Nielsen J, Lachapelle G. Gps vulnerability to spoofing threats and a review of antispoofing techniques. Int J Navigation Observation 2012. http://dx.doi.org/10.1155/2012/127072.

Likert R. A technique for the measurement of attitudes. Archives Psychol 1932;22(140).

Manadhata P, Wing J. An attack surface metric. IEEE Trans Softw Eng 2011;37(3):371–86. http://dx.doi.org/10.1109/TSE.2010.60.

McGraw G. Software security: building security in. Addison-Wesley software security series. Addison-Wesley; 2006.

Messerges TS, Dabbish EA, Sloan RH. Examining smart-card security under the threat of power analysis attacks. IEEE Trans Comput 2002;51(5):541–52. http://dx.doi.org/10.1109/

TC.2002.1004593. URL, http://dx.doi.org/10.1109/TC.2002.1004593.
Miller C, Valasek C. Adventures in automotive networks and control units. Accessed on March 2014. 2014. URL, http://blog.ioactive.com/2013/08/car-hacking-content.html.
National Computer Security Center (NCSC). Glossary of computer security terms. Tech. Rep. NSCD-TG-004. 1988. Fort Meade, Md.
OWASP. Risk rating methodology. December 2011. URL, https://www.owasp.org/index.php/OWASP_Risk_Rating_Methodology#Approach.
Pardue H, Landry J, Yasinsac A. A risk assessment model for voting systems using threat trees and monte carlo simulation. In: 2009 First International Workshop on Requirements Engineering for e-Voting Systems (RE-VOTE), Atlanta, GA; 2009. p. 55–60. http://dx.doi.org/10.1109/RE-VOTE.2009.1.
Perlroth N. Cameras may open up the board room to hackers. January 2012. URL, http://www.nytimes.com/2012/01/23/technology/flaws-in-videoconferencing-systems-put-boardrooms-at-risk.html?pagewanted=all.
Physical Measurement Laboratory of NIST. The nist reference on constant, units, and uncertainty–uncertainty of measurement results. Oct. 2000. URL, http://physics.nist.gov/cuu/Uncertainty/international1.html.
Risk Steering Committee. DHS risk lexicon, u.S. Department of Homeland Security. Sep. 2008. URL, https://www.dhs.gov/xlibrary/assets/dhs_risk_lexicon.pdf.
Ruddle A. Security risk analysis approach for on-board vehicle networks. In: The fully networked car workshop at the Geneva international motor show; March 2010. URL, http://evita-project.org/Publications/Rud10.pdf.
Shafer G. A mathematical theory of evidence. Princeton, NJ: Princeton University Press; 1976.
Shirey R. Internet security glossary, version 2. RFC 4949 (Informational). Aug 2007. URL, http://www.ietf.org/rfc/rfc4949.txt.
Stoneburner G, Goguen A, Feringa A. Risk management guide for information technology systems. NIST Special Publication, Vol. 800(30). Gaithersburg, Md: U.S. Dept. of Commerce, National Institute of Standards and Technology; 2002.
The TREsPASS Project. Technology-supported risk estimation by predictive assessment of socio-technical security. Oct. 2014. URL, http://www.trespass-project.eu/.
Wheeler E. Security risk management: building an information security risk management program from the ground up. Waltham, MA: Elsevier Science; 2011. p. 105–25. Ch. Risk Exposure Factors.
Wohlin C, Runeson P, Host M, Ohlsson M, Regnell B, Wesslen A. Experimentation in software engineering. Berlin Heidelberg: Springer-Verlag; 2012.
Working Group 1 of the Joint Committee for Guides in Metrology (JCGM/WG 1). International vocabulary of metrology – basic and general concepts and associated terms (vim). 3rd ed. 2012 URL http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf.
Yasinsac A. Insider threats to voting systems. In: Proc.of the 2010 workshop on governance of technology, information and policies, GTIP '10; 2010. p. 1–8. http://dx.doi.org/10.1145/1920320.1920321.
Zadeh LA. Fuzzy logic and approximate reasoning. Synthese 1975;30:407–28. http://dx.doi.org/10.1007/BF00485052. URL, http://dx.doi.org/10.1007/BF00485052.

Lotfi ben Othmane is currently a Scientific Researcher at Fraunhofer SIT, Germany. Previously, he worked at Lero-The Irish Software Engineering Research Center, Ireland and at the Eindhoven University of Technology, The Netherlands, as a post-doctoral researcher. He received his Ph.D. degree from Western Michigan University (WMU), USA, in 2010, M.S. degree in Computer Science from University of Sherbrooke, Canada, in 2000, and B.S degree from University of Sfax, Tunisia, in 1995. He is currently investigating the development of secure evolving software.

Rohit Ranchal is a PhD Candidate in Computer Science at Purdue University. He received his Ph. D. in Computer Science from Purdue University in 2014, M.S. in Computer Science from Purdue University in 2011, and B.Tech. in Information Technology from DAVIET Jalandhar, India in 2009. His research interests include monitoring, anomaly detection, failure diagnosis, policy enforcement, SLA compliance and access control in service computing systems particularly cloud computing and SOA environments. He has received ACM Graduate Teaching Assistant Award in 2013, Raymond Boyce Graduate Teacher Award and Teaching Academy Graduate Teaching Award in 2014. He is a member of IEEE.

Ruchith Fernando is currently with Amazon.com, Inc. He received his Ph.D. from Purdue University. He contributed to several research projects and has about 10 peer-reviewed publications. He received the Honorable Mention in the Research Poster Competition of 13th Annual Information Security Symposium, CERIAS. His main research interest include: user-centric identity management and security for web services.

Bharat Bhargava is currently a Professor of computer science at Purdue University. He received the BE degree from the Indian Institute of Science, and the MS and PhD degrees in electrical engineering from Purdue University, West Lafayette, IN. He His research involves mobile wireless networks, secure routing and dealing with malicious hosts, providing security in Service Oriented Architectures, adapting to attacks, and experimental studies. His name has been included in the Book of Great Teachers at Purdue University. Moreover, he was selected by the student chapter of ACM at Purdue University for the Best Teacher Award. He is a fellow of the IEEE.

Eric Bodden is heading the Secure Software Engineering group at Fraunhofer SIT and Technische Universität Darmstadt. He received his Ph.D. from McGill University. His research aims at aiding software engineers to build more secure software. This includes the design and development of novel algorithms, methods and tools for static and dynamic program analysis but also research on the proper development processes that safeguard against the introduction of software vulnerabilities. His group collaborates with various renowned research institutes and Fortune 500 companies.