Trento 2013 Security Engineering Course: Experiment Description

April 17, 2014

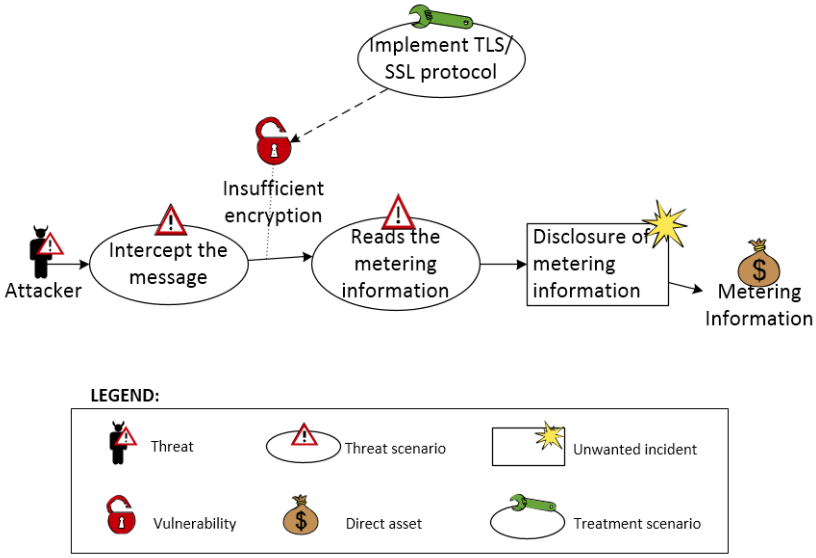
1 Research method

This section describes the design of the performed experiment, following the guidelines by Wohlin et al. [1].

1.1 Selection of methods

CORAS is a visual method which consists of three tightly integrated parts, namely, a method for risk analysis, a language for risk modeling, and a tool to support the risk analysis process. The risk analysis in CORAS is a structured and systematic process which use diagrams (see Figure 1(a)) to document the result of the execution of each step. The steps are based on the international standard ISO 31000 [2] for risk management: context establishment, risk analysis (that identifies assets, unwanted incidents, threats and vulnerabilities), and risk treatments.

The Security Requirements Engineering Process (SREP) is an asset-based and risk-driven method for the establishment of security requirements in the development of secure Information Systems. SREP supports a micro-process, consisting of nine steps: agree on definitions, identify critical assets, identify security objectives, identify threats and develop artifacts, risk assessment, elicit security requirements, categorize and prioritize security requirements, requirements inspection, and repository improvement. The result of the execution of each step of the process is represented using tables or natural language (see Figure 1(b)). SREP is compliant with international standards



(a) CORAS - Threat Diagram

Name of Misuse Case: Spoof of information		
ID 1		
Summary: the attacker gains access to the message exchange between the SM and SNN and disclose the secret exchange of information		
Probability: Frequent		
Preconditions: 1) The attacker have access to the communication channel between SM and SNN		
User Interactions	Misuser interactions	System Interaction
The SM sends the information about power consumption		
	The attacker reads the information	
		The SSN receives the information without knowing that someone have read the message
Postconditions: 1) The attacker knows personal information about the power consumption of the customer		

(b) SREP - Threat Specification using misuse cases

Figure 1: Examples of Visual (CORAS) and Textual (SREP) Methods' Artefacts.

ISO/IEC 27002 [3] and ISO/IEC 15408 [4] within the scope of requirements engineering and security management.

For additional details about CORAS and SREP we refer the reader to [5, Chap. 3] and [6]. Note that, in the rest of the paper, we denote with "security requirements" both the concepts "treatments" in CORAS and "security requirements" in SREP because they have the same semantic: they are both defined as a means to reduce the risk level associated with a threat.

1.2 Research approach

The *goal* of the experiment was to evaluate and compare two types of risk-driven methods, namely, visual methods (CORAS) and textual methods (SREP) with respect to their *effectiveness* in identifying threats and security requirements, and the participants' *perception* of the two methods. Hence, visual and textual methods were the two treatments that we have considered in the experiment. We want to investigate the following research questions:

RQ1 *Is the effectiveness of the methods significantly different between the two type of methods?*

RQ2 *Does the effectiveness of the methods vary with the assigned tasks?*

RQ3 *Is the participants' preference of the method significantly different between the two type of methods?*

RQ4 *Is the participants' perceived ease of use of the method significantly different between the two type of methods?*

RQ5 *Is the participants' perceived usefulness of the method significantly different between the two type of methods?*

RQ6 *Is the participants' intention to use the method significantly different between the two type of methods?*

To answer the first two research questions we have measured *effectiveness* by counting the number of threats and the number of security requirements as the main outcomes of the methods' application (as done in [7,8]). Research questions $RQ3 - RQ6$ have been answered by measuring perception-based variables *perceived usefulness* (PU), *perceived ease of use* (PEOU), *intention*

to use (ITU) with a post-task questionnaire. In order to gain a better understanding of *why a method is effective* (or more effective than another) we also carried out individual interviews with the participants.

1.3 Hypotheses

We have translated research questions $RQ1 - RQ6$ into a list of null hypotheses to be statistically tested. Due to the lack of space we report here only the main alternative hypotheses to the null ones denoted as Hn_A where n specifies the research question to which the hypothesis is related and the index A specifies that is an alternative hypothesis.

$H1.1_A$ There will be a difference in the number of threats found with the visual method and with the textual method

$H1.2_A$ There will be a difference in the number of security requirements found with the visual method and with the textual method

$H2.1_A$ There will be a difference in the number of threats found with the visual and the textual method within each facet

$H2.2_A$ There will be a difference in the number of security requirements found with the visual and the textual method within each facet

$H3_A$ There will be a difference in the participants preference for the visual and the textual method

$H4_A$ There will be a difference in the participants perceived ease of use for the visual and the textual method

$H5_A$ There will be a difference in the participants perceived usefulness for the visual and the textual method

$H6_A$ There will be a difference in the participants intention to use for the visual and the textual method

Hypotheses $H1.1_A$-$H1.2_A$ are related to $RQ1$ and suppose that there will be a difference in the effectiveness of the methods. $H2.1_A$-$H2.2_A$ assume a possible relation between the effectiveness of the methods and the facets on which the methods is applied (RQ2). Hypothesis $H3_A$ assumes there will be a difference in the participants' overall preference for the methods (RQ3).

$H4_A$-$H6_A$ assume that the participants' perceived easy of use, perceived usefulness, and intention to use variables will differ for the two methods (RQ4-RQ6).

1.4 Experimental design

Participants for the experiments were recruited among master students enrolled in the Security Engineering course at the University of Trento. The participants had no previous knowledge of the methods under evaluation. A within-subject design where all participants apply both methods was chosen to ensure a sufficient number of observations to produce significant conclusions. In order to avoid learning effects, the participants had to identify threats and mitigations for different types of security facets of a Smart Grid application scenario. The Smart Grid is an electricity network that can integrate in a cost-efficient manner the behavior and actions of all users connected to it like generators, and consumers. They use information and communication technologies to optimize the transmission and distribution of electricity from suppliers to consumers.

The tasks differ in the security facets for which the groups had to identify threats and security requirements. The security facets included Security Management (Mgmnt), Application/Database Security (App/DB), Network/Telecommunication Security (Net/Teleco), and Mobile Security (Mobile). For example, in the App/DB facet, groups had to identify application and database security threats like cross-site scripting or aggregation attacks and propose mitigations.

The participants were divided into 16 groups so that each group would apply the visual method (CORAS) to exactly two facets and the textual method (SREP) to the remaining two facets. For each facet, the method to be applied by the groups was randomly determined. Table 1 shows for each facet the number of groups assigned to visual and textual methods.

1.5 Experimental Procedure

The experiment was performed during the Security Engineering course held at University of Trento from September 2012 to January 2013. The experiment was organized in three main phases:

Facet/Method	Visual	Textual
Mgmnt	6	10
App/DB	9	7
Net/Teleco	9	7
Mobile	8	8

Table 1: Experimental design

- **Training**. Participants were given a tutorial on the Smart Grid application scenario and a tutorial on visual and textual methods of the duration of two hours each. The Smart Grid scenario focused on the gathering of metering information from the smart meters and their transmission to the utility services for billing purposes. Then, participants were administered a questionnaire to collect information about their background and their previous knowledge of other methods and they were divided into groups based on the experimental design.

- **Application**. Once trained on the Smart Grid scenario and the methods, the groups had to repeat the application of the methods on four different facets: Security Management, Application/Database Security, Network Security and Mobile Security. For each facet, the groups:

  - Attended a two hours lecture on the threats and possible mitigations specific for the facet but not concretely applied to the case study.
  - Had one week to apply the assigned method to identify threats and security requirements specific for the facet.
  - Gave a short presentation about the preliminary results of the method application and received feedback.
  - Had one week to deliver an intermediate report to get feedback.

  At the end of the course in mid January 2013, each group submitted a final report documenting the application of the methods on the four facets.

- **Evaluation**. In this phase, the experimenters (the authors of this paper) assessed participants final reports while the participants evaluated the method through questionnaires and interviews. First, each group gave a presentation summarizing their work in front of the experimenters and of the expert. The expert evaluated the quality of the threats and the mitigations proposed for the Smart Grid application scenario. Then, participants were administered the post-task questionnaire to be filled in online. Last, each participant was interviewed for half an hour by one of the experimenters to investigate which are the advantages and disadvantages of the methods.

The interview guide contained open questions about the overall opinion of the methods, their advantages and disadvantages, the difficulties encountered during the application of the methods and the main differences among them. The interview questions were the same for all the interviewees even though some specific questions were added for some of the participants when their answers to the questionnaire were contradictory. The questions are reported in Table 2 in Appendix.

The questionnaire was adapted from the questionnaire reported in [7] which was inspired to the Technology Acceptance Model (TAM) [9]. The questionnaire consisted of 22 questions which were formulated in an opposite statements (positive statement on the right and negative statement on the left) format with answers on a 5-point Likert scale. The questions were formulated as follows: Q1: Whether the method was easy or hard to use; Q2: The method made the security analysis easier or harder than an ad hoc approach; Q3: The method was easy or difficult to master; Q4: Intention to use the method to identify threats and security requirements in a future project course; Q5: The method is better in identifying threats and security requirements than using common sense; Q6: Intention to use the method to identify threats and security requirements in a future project at work; Q7: Confusion about how to apply the method to the problem; Q8: Whether the method made the search for threats and security requirements more or less systematic; Q9: Intention to use the method if suggested by someone at work; Q10: Whether the method would be easy or hard to remember; Q11: Whether the method makes more or less productive in identifying threats and security requirements; Q12: Intention to use the method in a discussion with a customer; Q13: Whether the process of the method is well or not well detailed; Q14-Q15: A catalog of threats and security requirements

makes easier or harder the security analysis with the method; Q16-Q17: The method helps or not helps in brainstorming on the threats and the security requirements; Q18: Whether the tool is easy or hard to use (asked just for the visual method because it had tool support); Q19-Q22: Difficulties of facets. To avoid that the participants answered on "auto-pilot", some of the questions (e.g. Q2, Q10, Q13) were given with the most positive response on the left and the most negative on the right.

References

[1] C. Wohlin, P. Runeson, M. Hst, M. C. Ohlsson, B. Regnell, and A. Wessln, *Experimentation in software engineering*. Springer, 2012.

[2] *ISO 31000 Risk management – Principles and guidelines*, International Organization for Standardization, 2009. [Online]. Available: http://en.wikipedia.org/wiki/ISO_31000

[3] *ISO/IEC 27002 Information technology - Security techniques - Code of practice for information security management*, International Organization for Standardization and International Electrotechnical Commission, 2005. [Online]. Available: http://en.wikipedia.org/wiki/ISO/IEC_27002

[4] *ISO/IEC 15408 Information technology Security techniques Evaluation criteria for IT security*, International Organization for Standardization and International Electrotechnical Commission, 2005. [Online]. Available: http://www.enisa.europa.eu/activities/risk-management/current-risk/laws-regulation/rm-ra-standards/iso-iec-standard-15408

[5] M. S. Lund, B. Solhaug, and K. Stølen, *Model-driven risk analysis: the CORAS approach*. Springer, 2011.

[6] D. Mellado, E. Fernández-Medina, and M. Piattini, "Applying a security requirements engineering process," in *Proc. of the 11th European Symposium on Research in Computer Security (ESORICS)*. Springer, 2006, pp. 192–206.

[7] A. L. Opdahl and G. Sindre, "Experimental comparison of attack trees and misuse cases for security threat identification," *Information and Software Technology*, vol. 51, no. 5, pp. 916–932, 2009.

[8] A. Teh, E. Baniassad, D. Van Rooy, and C. Boughton, "Social psychology and software teams: Establishing task-effective group norms," *Software, IEEE*, vol. 29, no. 4, pp. 53–58, 2012.

[9] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, pp. 319–340, 1989.

Interview Questions
What do you think about method?
Do you think the method is an easy method to apply? Why?
While applying the method where you got confused about how to apply it?
Do you think the method helps you brainstorming? Why?
Do you think the method helped you to identify threats and security requirements?
Which are the advantages of the method?
Which are the disadvantages of the method?
Would you use the method in the future?
What do you think about CORAS tool?
Do you think CORAS tool is hard to use? Why?
Which version of the CORAS tool did you use?
Which do you think are the significant differences between the two methods?
Which was according to you the most difficult facet? And why?

Note: These questions were asked both for the visual (CORAS) and the textual method (SREP).

Table 2: Interview Guide