Towards a Common Evaluation Framework for Cyber Security Visualizations

Noëlle Rakotondravony[1]
nr@sec.uni-passau.de
[1]University of Passau, Germany

Hans P. Reiser[1,2]
hr@sec.uni-passau.de
[2]Karlsruhe Institute of Technology, Germany

1 Introduction

With the increasing number of cyber security visualization tools, literature has repeatedly highlighted the need for a common evaluation framework to help assessing and quantifying the effectiveness of proposed tools and validate their adequacy to targeted usages. However, there is no research yet that supports the development of such framework. We present a work in progress and preliminary ideas towards building a common evaluation framework for cyber security visualizations.

Many authors have contributed to the state of the art of evaluation in the field of cyber security visualization. Staheli et al. [4] systematized a taxonomy for evaluable components, which have also been derived from evaluation work in other research fields. The methodologies described by Sethi et al. [3] and Suo et al. [5] reflect the different forms of an evaluation framework for cyber security visualization and methods to address its development.

Moreover, through an analysis of research works from the VizSec venue, Staheli et al. [4] highlighted that among the forms of evaluation that are included in the selected works, some dimensions and evaluation techniques (e.g., psychophysiological methods) are not yet considered. This is especially because doing so would require knowledge that is not necessarily covered by the expertise of cyber security analysts. Besides, considering a target user's requirements as evaluation metrics, as described by Sethi et al. [3], is undoubtedly an important form of evaluation but covers only qualitative aspects. In order to build a comprehensive framework, this could also be completed with other relevant aspects such as quantitative ones.

1.1 Motivation and objectives

Existing research works in literature show that the evaluation cannot be exclusively studied from a security experts point of view, nor by excluding research findings from other relevant domains. Visualization for cyber security is at the intersection of parent fields such as computer security, information visualization, human-computer interaction, and software design. Therefore, applying and adapting the established knowledge on evaluation processes from these fields constitute a promising route. A challenge when building a comprehensive framework for evaluation in cyber security visualization is to find clear and adequate directives on how to organize external findings so that they effectively contribute to enrich the cyber security visualization field itself. However, no proposition has been made yet on how such framework should be defined and how its different components could be organized to make it usable for the community.

Overall, we aim to develop a common framework as a set of guidelines that an evaluator can follow when conducting an evaluation of cyber security visualizations. At this stage of our analysis, we identified the following objectives to be relevant contributions:

- Develop a framework as a set of reference guidelines to follow in evaluation processes. The framework should provide the evaluator all necessary means to quantify the effectiveness of considered tools according to his/her objectives and roles, taking into consideration the evaluation aspects or purposes (summative: when the goal is to understand the quality of a tool, formative: when the goal is to learn how to design better tools), the tool usage purposes, the targeted users and other relevant configurations.

- Develop a framework which takes into consideration the evaluation methods and techniques from parent fields, reasonably and scientifically combines them with cyber security field-related methods or requirements in order to provide an applicable comprehensive evaluation of the effectiveness of cyber security visualization tools.

- Develop reusable or repeatable evaluation guidelines. The goal is to address the fact that a changing nature of threat subjects usually leads to a change of security-related tasks, making it difficult to reuse previous evaluation or compare existing tools against newer ones.

1.2 Relevance to community

The relevance of having an evaluation framework for cyber security visualizations has been formulated already from the early age of the field. A set of clear theoretical guidelines and methodological principles would allow quantifying the effectiveness of visualization systems [2]. Consequently, it constitutes a relevant support for the adoption of the tools or the design techniques by users.

Moreover, having a common evaluation framework available in literature would also aid in the research and review processes, allowing to continuously produce quality papers, that are fairly judged based on detailed metrics that allow reviewers to expect certain results when considering defined evaluation techniques.

2 Methodology

So far, our methodology to construct a common, comprehensive and practicable evaluation framework for cyber security visualizations comprises several major steps whose goals and subgoals are described in this section.

2.1 Define the different aspects of evaluation in visualization fields

In this step, we want to systematically study the *evaluation techniques* in visualization fields. The aspects to be considered include (but are not limited to):

- The possible forms, purposes of an evaluation task and the corresponding applicable methodologies available in literature.
- The dimensions, components to evaluate, as initiated by Staheli et al. [4], with a detailed perspective on the different variables or parameters that compose them, and that are actually measured to objectively quantify the effectiveness of a tool.

- The factors influencing the set of variables and metrics to be considered in an evaluation. These include: user's expertise and awareness (for example, whether the user knows or not what he is looking for when using the tool), the expertise of the evaluator (for example, we could assume that one cannot properly evaluate without a background in any of the parent fields), user's working environment, etc.

Through a literature review, interviews with visualization users, an eventual collaboration with experts, we aim to provide a concise description of what constitutes an evaluation task, find adequate terminologies and highlight the research fields involved in the construction of a evaluation comprehension framework.

2.2 Define the scope of the evaluation framework

In this step, we want to study the particular scope of an evaluation framework that makes it specific for cyber security visualization field. The goal is to find the elements that should be considered by an evaluator when investigating the tools following a selected configuration of the evaluation tasks. We propose to conduct our analysis from two perspectives: from practice point of view, with the objective to discover proper practical characteristics of analysts' (or target users') activities when using visualization tools, and from literature point of view, with the objective to derive theses characteristics from research studies that analyzed the tasks of security analysts using, for example, Cognitive Task Analysis (CTA) methodology [1].

The utility of having such highlights in the framework is that whenever the evaluation task does not consider cyber security field-related metrics, already established techniques from parent fields are applicable. Otherwise, a challenge is to find the correct methods to combine such techniques or adapt them in a way that allows the evaluator to consider proper security metrics in his measurements.

2.3 Formulate the "evaluation questions" and answers

In this step, we want to make the evaluation framework more usable for evaluators. Our idea consists of formulating the evaluator's tasks objectives into different questions, called "*evaluation questions*". Examples are *Did the tool allow the user gain all necessary knowledge to achieve his task?* or *How fast could the user reach his goal using the tool?*. Trying to answer the *evaluation questions* will guide the evaluator into finding the set of variables that are interesting to measure, e.g., *the list of information communicated by the graphs* for the first evaluation question. In fact, identifying the questions associated to the selected evaluation's purposes helps to define:

- Which qualities are expected from the evaluated visualization system to answer the evaluation question?
- What are the different variables that the evaluator can quantitatively measure or qualitatively appreciate to decide on the effectiveness of proposed tools?
- Which field knowledge is to be considered and which techniques apply?

Formulating the *evaluation questions* ensures the flexibility of the framework and its reusability to various use-cases. In fact, evaluators' objectives might differ in some aspects but might also have similarities. In all cases, the guidelines in the framework should allow setting similar expectations for evaluation questions that require and appy similar evaluation methodologies.

Furthermore, we propose to gather the *evaluation questions* into a repository, which will be updated through a collaboration between practitioners and experts from different fields relevant to cyber security visualization. The overall architecture of such repository is yet to be defined. But in general, the idea is that it serves as reference location of elements to guide evaluators into asking the right questions (with corresponding answers) that help them set up the adequate evaluation methodology.

2.4 Translate the evaluation framework into practice

This step essentially focuses on helping users adopt the evaluation framework, for example through an education of users on the use of the evaluation in different interesting use-cases such as formative or summative evaluation. The distinctions between the two are yet hard to describe, but ideally, the guidelines in the evaluation framework are applicable in the evaluation and comparison of cyber security visualization tools as well as in the design of more effective ones.

3 Current project status

Currently, we are working on the step described in Section 2.1. Through a systematic literature review, we aim at progressively developing an exhaustive summary of existing literature. In parallel, at this early stage of our idea, we work on an iterative refinement of the proposed method, to get sharper formulation of the goals and contributions, and especially to shape the evaluation methodologies to later measure the qualities (utility, practicability,...) of the targeted evaluation framework.

To evaluate the framework, we want to make our test use-cases as diverse as possible to reflect the plurality of real-life scenarios in which the evaluation framework can be used. Use-cases include so far an *evaluation of the repeatability of the framework's guidelines*: for the same visualization tool, the same evaluation guidelines are applied by separate groups of evaluators and a similar or comparable output is expected. This also allows to assess the reproducibility of the metrics and methodologies. Another use-case consists of *comparing two or more visualization tools*: with the same evaluation configurations and methodologies, similar or comparable outputs are expected from different groups of evaluators.

4 Conclusion

We described our idea on how to structure and develop a common evaluation framework for cyber security visualizations, whose relevance has been repeatedly identified in the field. Building such evaluation framework is not an easy task. Particularly, we believe that the evaluation framework should be not only comprehensive and repeatable, but also as dynamic as the evolution of techniques and technologies used to build the tools. The methodology which we propose describes the steps towards building such evaluation framework, yet remains open to discussions on how to improve the quality of the targeted outcome and ensure its adequacy to the community.

Acknowledgments

The research leading to these results was supported by the "Bavarian State Ministry of Education, Science and the Arts" as part of the FORSEC research association.

References

[1] A. D'Amico and K. Whitley. The real work of computer network defense analysts. In *Proceedings of the Workshop on Visualization for Computer Security*, VizSEC '07, pp. 19–37. Springer, 2008.

[2] P. Ren. Ensuring the continuing success of VizSec. In *Proceedings of the 3rd International Workshop on Visualization for Computer Security*, VizSEC '06, pp. 67–70. ACM, New York, NY, USA, 2006.

[3] A. Sethi, F. Paci, and G. Wills. EEVi–framework and guidelines to evaluate the effectiveness of cyber-security visualization. *The International Journal of Intelligent Computing Research (IJICR)*, 7(4):761–770, 2016.

[4] D. Staheli, T. Yu, R. J. Crouser, S. Damodaran, K. Nam, D. O'Gwynn, S. McKenna, and L. Harrison. Visualization evaluation for cyber security: Trends and future directions. In *Proceedings of the 11th Workshop on Visualization for Cyber Security*, VizSec '14, pp. 49–56. ACM, 2014.

[5] X. Suo, Y. Zhu, and G. S. Owen. Measuring the complexity of computer security visualization designs. In *Proceedings of the Workshop on Visualization for Cyber Security*, VizSec '08, pp. 53–66. Springer, 2008.