Security Pattern Mining: Systematic Review and Proposal

Santiago Moral-García[1], Santiago Moral-Rubio[2] and Eduardo Fernández-Medina[3]

[1] Kybele Group, Dep. of Computer Languages and Systems II, University Rey Juan Carlos
Tulipán s/n, 28933, Móstoles, Madrid, Spain

[2] Dep. Logical Security, BBVA, Batanes 3, 28760, Tres Cantos, Madrid, Spain

[3] GSyA Research Group, Dep. of Information Technologies and Systems
University of Castilla-la Mancha, Paseo de la Universidad, 4, Ciudad Real, Spain

Abstract. Organizations have suffered an increase in cyber attacks in recent years. For this reason, they need to guarantee confidentiality, integrity and availability of their information assets. To do this, they should seek support from security architectures. Security patterns are a good way to design security architectures, but most current security patterns are not applicable to this field. In a previous work, we defined a new pattern template to support the design of security architectures. After that work, we realized that it was necessary to discover and identify new security patterns adapted to this template, in order to facilitate the work of those security engineers who design architectures. A good way to discover and identify new patterns is pattern mining; therefore, in this paper we have carried out a Systematic Review (SR) of security pattern mining. After performing the SR, we have reached the conclusion that the proposals analyzed do not fulfill all main requirements to cover our needs. That's the reason why we have defined a high-level architecture of a new framework to discover, design and document security patterns focused on the design of security architectures.

1 Introduction

Information Security is currently an important discipline in the management of large organizations, due to the increase in cyber attacks that they have suffered in recent years [7]. For this reason, the main need of an organization, in terms of security, is to guarantee the fundamental security properties of their information assets, understanding security properties as confidentiality, integrity and availability. To do this, organizations should seek support from security architectures in order to mitigate the impact of the constant evolution of the organization's settings [11], and thus minimize the risks associated with each of them.

The concept of security architecture can be defined as a complete, structured, coordinated and rigorous design of information systems that supports a business

Moral-García S., Moral-Rubio S. and Fernández-Medina E..
Security Pattern Mining: Systematic Review and Proposal.
DOI: 10.5220/0003558900130024
In *Proceedings of the 8th International Workshop on Security in Information Systems* (WOSIS-2011), pages 13-24
ISBN: 978-989-8425-61-4

process in order to reduce the risk of confidentiality, integrity and availability when managing its information assets [16]. Moreover, Security architectures must identify the security mechanisms that are needed to define the system's security.

The security mechanisms used in security architectures are artifacts designed to prevent, detect and respond to information security incidents, in order to manage and reduce the confidentiality, integrity and availability of business processes information risks. A security mechanism cannot be used in isolation to protect a business process, however, a wide set of security mechanisms can reduce security risks when managing information assets in a business process. Security architecture therefore consists of a wide set of security mechanisms, which is complete, structured, coordinated and rigorous [24].

A good way to design security architectures is the use of security patterns, since they provide solutions to recurring problems related to systems' security [23]. Another advantage of security patterns is that they include extensive, accumulated and structured knowledge about security, thus providing guidelines for the construction and evaluation of secure systems [4].

Previous to this work, we carried out a Systematic Review (SR) of security patterns and we checked that the number of published security patterns and classifications of them has increased considerably in recent years [5, 6, 14, 19, 23, 25]. In that work, we discovered that the vast majority of current security patterns are focused on supporting the construction of new security mechanisms. In other work [15], we also discovered that most current security patterns are not applicable to the design of security architectures, since they do not consider organizations' settings, i.e. the variation between people, technologies, risks, processes, volumes of information, business strategies, etc. For this reason, we defined a new pattern template with the aim of supporting the design of security architectures [13].

Once the template had been developed, we realized that it was necessary to discover and document security patterns adapted to the new pattern template, in order to facilitate the work of organizations' security engineers when designing new architectures. Pattern mining is a good way to cover this need, since it is a good discipline to discover and identify patterns [18]. In this paper, we have therefore carried out a SR of security pattern mining in order to identify sources, frameworks, methods or methodologies that help us to discover and identify recurring problems related to systems' security. In order to ensure that this SR was performed in a structured and organized manner, we have used the guidelines proposed by [12 and the review protocol template developed by [2].

After the accomplishment of the SR, we have reached the conclusion that the initiatives analyzed do not fulfill all main requirements that are necessary to discover and document new security patterns adapted to the new pattern template developed. For this reason, we have introduced a high-level architecture of a new framework based on pattern mining for discovering, designing and documenting security patterns focused on supporting the design of security architectures.

The remainder of this paper is organized as follows. Section 2 provides a brief description of pattern mining. In section 3, we carry out an SR of security pattern mining. Section 4 defines a new framework based on security pattern mining. We finish with some conclusions and future work in section 5.

2 Pattern Mining

The concept of pattern mining can be defined as the process or technique of identifying proven solutions [22]. Although many patterns are already available for different domains, mining new patterns will remain an important activity for the future [3].

Norman Kerth and Ward Cunningham considered three general approaches to pattern mining [10]: the *introspective*, the *artifactual* and the *sociological* approach. The introspective approach is when researchers analyze systems that they have built and try to discover patterns basing on their experience. This approach allows us to discover patterns through individual experiences. The artifactual approach is when researchers analyze systems that have been built by different people who were solving a similar problem. This approach allows us to find points in common between different systems analyzed. The sociological approach is when researchers ask people that have built similar systems how they solved particular problems. This approach allows us to find recurring problems in the systems analyzed.

However, Linda Rising and David DeLano broke down pattern mining activities into three categories [18]: *individual contributions*, *second-hand contributions* and *patterns mining workshops*. In the first category, individual contributions, researchers write a pattern from his or her own experience. This category is similar to the introspective approach previously described. In the second category, secondhand contributions, researchers obtain or find information from interviews with experts or from the experience of others. This category is a union of the artifactual and sociological approach previously described. In the third category, patterns mining workshops, a group of researchers work together discussing about the new patterns that they find out within a specific scope.

After analyzing the approaches and categories of pattern mining shown above, we have observed that the artifactual approach, the sociological approach and the category of patterns mining workshop can be adapted when performing security pattern mining. The only difference is that researchers should analyze information security documents, instead of analyzing systems.

3 A Systematic Review of Security Pattern Mining

In this section, we have carried out a systematic review (SR) of security pattern mining. We have followed the review protocol template developed by Biolchini et al. [2] in order to facilitate SR planning and execution. The phases included in this review are the following: *review planning*, *review method* and *results analysis*.

3.1 Review Planning

In this phase, we have defined the SR objectives and we have shown the way to execute the review. To do this, we have defined the formulation of the research question and we have planned how the selection of sources and studies will be carried out. The sections included in this phase are the following: *question formularization*, *sources selection* and *studies selection*.

3.1.1 Question Formularization

The *question focus* of this SR is to identify the most relevant initiatives, studies and reports in the Information Security Engineering discipline which use security pattern mining to discover and identify solutions for recurring security problems.

The biggest *problem* that we have encountered in the design of security architectures is that several information security engineers provide different solutions to the same problem, because each one has his or her own guidelines or follows his/her own criteria. Noting the large number of security engineers who can work for the same organization, it is necessary to find an optimal and proven set of security guidelines to assist engineers when designing secure architectures. In recent years, many authors have described security solutions, but most of the repositories based on security patterns are not applicable to the design of security architectures in real environments [15.

The *research question* that will be addressed in this research is the following: *Which studies have been carried out to discover and identify solutions for recurring security problems by means of Security Pattern Mining?* The *keywords* that cover this question and which will be used during the SR are: *patterns* and *anti-patterns* linked with *mining*. In addition, we have added a list of *related concepts* to information security to complete our study: *security*, *confidentiality*, *integrity*, *availability*, *identification*, *access control*, *privacy* and *auditability*. All of this will enable us to perform a more focused investigation on the field of security patterns mining.

The *population groups* that will be observed are publications in the selected data sources which use security patterns mining to discover and identify solutions for recurring security threats.

The *expected result* at the end of this SR is the identification of studies related to security pattern mining. If the number of identified studies is not enough, we will reach the conclusion that security pattern mining is a field that has not been investigated in depth.

The main *application areas* that will benefit from the SR results are information security engineers, information security areas, non-experts in the field of security, academics and researchers. Finally, we shall obtain a vision of the current situation in this field, which will allow us to detect shortcomings and new needs for security pattern mining.

3.1.2 Sources Selection

Before beginning the selection of sources where to execute the searches, we have carried out a primary search to obtain an initial idea and to evaluate the volume of studies that exist in relation to security pattern mining. We have also verified whether a SR in this field currently exists.

Information security is an area that is being published both in prestigious journals and a variety of conferences, but many of these conferences are not easily accessible and the search of all proposals would be an impossible task. For this reason, we have focused on publications of higher quality and demand sources, trusting that the proposals published in other forums, if they are really successful will be published in higher quality sources.

The *selection criteria* used to evaluate the sources are, therefore, the following: digital libraries which are available for all people via the Internet, digital libraries that can be accessed through the website of the Universidad Rey Juan Carlos and books focused on security patterns. In addition, all digital libraries must possess advanced search mechanisms, that is, searches oriented to the title, abstract, keywords or authors of the initiatives. Finally, the main sources of the *initial sources list*, over which the SR execution will be run, are as follows: *ACM digital library*, *IEEE digital library*, *SpringerLink*, *Science@Direct* and some *books* focused on *security patterns*.

3.1.3 Studies Selection

Having defined the sources selection, it is now necessary to define the studies selection. To do this, we have described the search chain, the inclusion and exclusion criteria and the evaluation criteria for the studies.

We have combined the selected keywords and related concepts with AND and OR connectors to obtain our *search chain*. Table 1 shows the general search chain. We have altered this search chain in order to adapt it to the syntax of specific search engines.

Table 1. General search chain.

(Mining **AND (**Pattern **OR** Antipattern**)** **AND (**Security **OR** Confidentiality **OR** Integrity **OR** Availability **OR** Identification **OR (**Access **AND** Control**) OR** Privacy **OR** Auditability**))**

Once the search chain has been defined, we have described the *inclusion and exclusion criteria*. On the one hand, all studies obtained after executing the search chain in the selected sources will be included in the review. On the other hand, all studies that, *first condition*, are not within the scope of the Computer Science discipline, *second condition*, are duplicated or, *third condition*, do not contain sources, frameworks, methods or methodologies to discover and identify solutions for recurring security problems by means of security pattern mining will be eliminated of the review.

These criteria have helped us to define the procedure by which the studies will be obtained and evaluated. The *procedure for the studies selection* consists of four steps. At the *first step*, the search chain will be executed in the selected sources and all studies returned by each of the sources will be collected. At the *second step*, all studies that are not within the scope of the Computer Science discipline or are duplicated will be eliminated. At the *third step*, the title, abstract and keywords of all remaining studies will be read and all studies that clearly meet the third condition of the exclusion criteria will be eliminated. At the *fourth step*, all remaining studies will be read fully and all studies that do not meet the third condition of the exclusion criteria will be selected as primary proposals.

3.2 Review Execution

After planning the SR, we have initiated the review execution. During this phase, we have evaluated the studies obtained following the procedure for studies selection previously defined. Finally, we have extracted from the selected studies the relevant information for the research question.

3.2.1 Selection Execution

The selection execution has been carried out following the four steps of the procedure for studies selection (see Section 3.1.3). After carrying out the *first step* of the procedure for studies selection, we have obtained *142 studies* in total. Once we have performed the *second step*, we have obtained *102 studies*. After carrying out the *third step*, we have obtained *8 studies*. Finally, once we have performed the *fourth step* of the procedure, we have obtained *2 primary studies*.

3.2.2 Information Extraction

After selecting the primary studies, we have extracted important information from them. To do this, we have analyzed in depth each of the initiatives and have performed a brief outline.

- *"A methodology for mining security tactics from security patterns"* [20]:

In this paper, authors introduce a methodology for discovering and identifying security tactics from well known security patterns. The aim of this methodology is to examine existing security patterns to verify if any of these patterns satisfies the conditions necessary to qualify it as a security tactic. These conditions include: atomicity, force limitation, problem-specificity, completeness and trade-offs between forces. After defining the methodology, they provide an example to demonstrate how the criteria described can be used to mine security tactics from existing security patterns. Finally, authors hope this research will begin to produce new architectures through tactics mining.

- *"Mining Security Patterns"*, Section 7.5 of [22]:

In section 7.5 of this book, Schumacher shows different approaches for mining security patterns. He basically considers two ways: on the one hand, he learns from typical errors and tries to discover how these errors could have been prevented, *artifactual approach*. On the other hand, he directly obtains information from security standards. In this section, Schumacher also introduces a wide list of information sources where security patterns can be discovered. This list includes *security information providers* (CERT, hackers groups, security companies, newsgroups, mailing lists, articles and books) and *security standards* (Common Criteria, IT Baseline Protection Manual, ISO 17779).

3.3 Results Analysis

After the SR execution, we have analyzed and summarized the primary studies obtained in the previous phase.

In relation to the study *"A methodology for mining security tactics from security patterns"* [20], we have observed that security tactics can be described in a hierarchy based on the following three categories: resisting attacks, detecting attacks and recovering from an attack. As authors say in the paper, this categorization can be useful at an early stage of development, but we have observed that security tactics are less useful when designing security architectures. This is because the hierarchy of categories should include subcategories related to the following questions: a) Which assets need to be protected? b) What are we protecting ourselves from? And c) Which security properties do we intend to conserve?

In relation to section 7.5, *"Mining Security Patterns"*, of the book [22], we have observed that both techniques and information sources shown are very useful to carry out current security patterns mining, but are not enough to discover, design and document security patterns oriented to the design of security architectures. This is due to the fact that security patterns focused on the design of security architectures include three abstraction levels in their solution and perform a qualitative analysis of the most important technological aspects with regard to the proposed solution [13]. Information sources provided by Schumacher could help us to discover, design and document the first solution level and in some cases the second level, but it is not common to find public security documents that include solutions oriented to technological products and detail the most important technological aspects.

We have therefore reached the conclusion that studies analized in this SR do not fulfill all main requirements for discovering, designing and documenting security patterns focused on the design of security patterns.

4 D^{3}SP: A New Framework based on Security Pattern Mining

In this section, we have introduced a high-level architecture of a new framework based on pattern mining for Discovering, Designing and Documenting Security Patterns (*D^{3}SP*). The main objective of D^{3}SP is to support the design of security architectures.

This framework contains: a set of sources for mining security patterns, a Pattern Template for Documenting Security Patterns (*PT-DSP*), a method for Discovering Security Patterns from Public Documents (*DSP-PD*), a method for Discovering Security Patterns from Deployed Architectures (*DSP-DA*), an Ontology of Security Patterns (*OSP*) and a method for Integrating Discovered Security Patterns (*IDSP*) into the ontology. Figure 1 shows each of the listed elements and how they compose the framework.

As can be observed in Figure 1, arrows represent methods and figures represent elements. Next, we have carried out a brief outline of each of the methods and elements that compose the framework D^{3}SP.

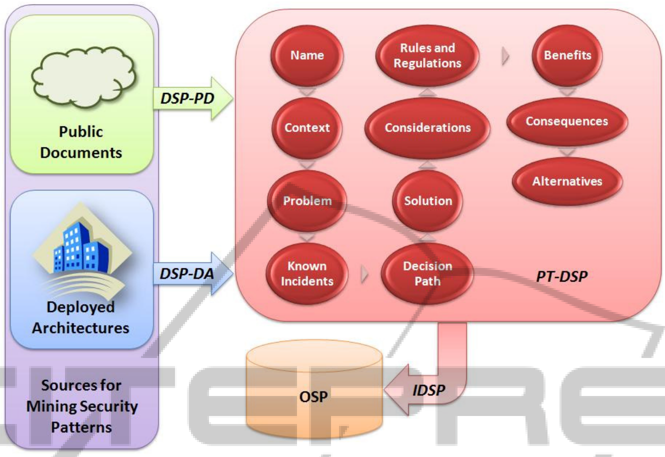

Fig. 1. Framework D^{3}SP.

4.1 Sources for Mining Security Patterns

We have selected two different types of sources to carry out security patterns mining. On the one hand, we have selected public sources which could be accessed by anyone via the Internet. On the other hand, we have selected private sources which could only be accessed if there is a direct contact with an organization. Public sources have been labeled as *public documents* and private sources are related to *deployed architectures* in large organizations.

4.1.1 Public Documents

We have defined public documents as reports, researches or papers related to systems' security which can be found in information sources hosted at websites.

After performing an initial search to find public documents which can help us to discover security patterns, we have selected the following information sources to carry out security pattern mining: the Open Web Application Security Project [17], SANS Institute [21], the Information Security Forum [9], the Information Systems Audit and Control Association [8] and the International Information Systems Security Certification Consortium [1]. We will perform all searches for public documents from this set of information sources.

4.1.2 Deployed Architectures

We have defined *deployed architectures* as designs of information systems which have been implemented in large organizations in order to reduce the risks of security when managing their information assets. These architectures must have been used for a significant time and have demonstrated to be effective for the organization in which they are implemented.

With the aim of achieving a wide range of deployed architectures, we have begun to maintain direct contact with a multinational entity of the financial sector.

4.2 A Pattern Template for Documenting Security Patterns (PT-DSP)

In a previous work [13], we proposed a new pattern template to describe and document security patterns focused on the design of security architectures. PT-DSP will help us to complement security patterns and make them more applicable to security architecture design environments. The elements included in the template are as follows: *Name*, *Context*, *Problem*, *Known Incidents*, *Decision Path*, *Solution*, *Considerations*, *Rules and Regulations*, *Benefits*, *Consequences* and *Alternatives*.

4.3 Method for Discovering Security Patterns from Public Documents (DSP-PD)

DSP-PD method is based on the *artifactual* approach of pattern mining proposed by [10]. After finding a set of public documents from any of the selected information sources, a researcher must analyze security solutions that have been built to solve a similar problem. The aim of this analysis is to find points in common between different solutions analyzed, in order to obtain an optimal solution to the given problem.

After carrying out the analysis, the researcher must adapt all relevant information found to design and document one or more security patterns following the elements included in PT-DSP.

4.4 Method for Discovering Security Patterns from Deployed Architectures (DSP-DA)

DSP-DA method is based on the artifactual and sociological approach of pattern mining proposed by [10]. After obtaining a set of deployed architectures from one or more large organizations, a researcher must, on the one hand, analyze designs of information systems that have been built by different people who were solving a similar problem, in order to find points in common between different designs (*artifactual approach*), and on the other hand, ask people that have designed similar information systems how they solved particular problems, in order to find recurring problems in the design analyzed (*sociological approach*).

After carrying out the two analyses, in the same way as in DSP-PD, the researcher must adapt all relevant information found to design and document one or more security patterns following the elements included in PT-DSP.

4.5 Ontology of Security Patterns (OSP)

OSP is a small ontology which contains definitions of security concepts (such as vulnerabilities, attacks, threats, security objectives, risk, types of information assets, etc.), elements included in PT-DSP (security context, security problem, decision path, security solution, etc.) and relationships between them. OSP provides a wide set of relationships between security concepts and security patterns, enabling security engineers to find useful information when designing security architectures. OSP should answer the following questions: a) What vulnerabilities are eliminated with the security solution "X"? b) What risks are associated with the vulnerability "Y"? c) What information assets are ensured with the security solution "Z"? d) etc.

4.6 Method for Integrating Discovered Security Patterns (IDSP)

IDSP method is based on the category of pattern mining, *patterns mining workshop*, proposed by [18]. After performing security pattern mining from public documents (DSP-PD method) or deployed architectures (DSP-DA method), some security patterns obtained can be invalid or can be a sub-pattern of a security pattern already found. For this reason, a group of researchers must work together to discuss about the new security patterns found. A new security pattern will only be integrated into the ontology if the group of researchers believes it is appropriate.

5 Conclusions and Future Work

Organizations should seek support from security architectures in order to minimize the risks associated to the constant evolution of the organization's settings. The problem is that several information security engineers provide different solutions when designing security architectures, because each one has his or her own guidelines. For this reason, it is necessary to provide security engineers with a wide set of security patterns focused on the design of security architectures.

With the aim of encountering techniques or processes to discover new security patterns, in this paper, we have carried out a Systematic Review (SR) of security pattern mining. The obtained result in this SR is that proposals analyzed do not fulfill the main requirements to carry out security pattern mining focused on the design of security architectures.

For this reason, after performing the SR, we have introduced a high-level architecture of a new framework based on security pattern mining. The main objective of this framework is to discover, design and document security patterns focused on supporting the design of security architectures. In future works, we have planned to formalize and develop each of the components included in the framework defined.

Acknowledgements

This research has been carried out in the framework of the following projects: MODEL-CAOS (TIN2008-03582/TIN) financed by the Spanish Ministry of Education and Science, BUSINESS (PET2008-0136) financed by the Ministry of Science and Innovation, and SISTEMAS (PII2I09-0150-3135) and SERENIDAD (PEII11-0327-7035), all financed by the Local Government of Castilla-La Mancha, in Spain.

References

1. (ISC)^2. (2011). The International Information Systems Security Certification Consortium, from http://www.isc2.org/
2. Biolchini, J., Mian, P. G., Natali, A. C. C. and Travassos, G. H. (2005). Systematic Review in Software Engineering. Systems Engineering and Computer Science Department COPPE / UFRJ: Rio de Janeiro.
3. Buschmann, F., Meunier, R., Rohnert, H., Sommerlad, P. and Stal., M. (1996). Pattern-oriented software architecture: A system of patterns. Wiley.
4. Fernandez, E., Washizaki, H., Yoshioka, N., Kubo, A. and Fukazawa, Y. (2008). Classifying Security Patterns Progress in WWW Research and Development (pp. 342-347).
5. Fernandez, E. B., Washizaki, H. and Yoshioka, N. (2008). Abstract security patterns. Proceedings of the 15th Conference on Pattern Languages of Programs, Nashville, Tennessee.
6. Hafiz, M., Adamczyk, P. and Johnson, R. E. (2007). Organizing Security Patterns. Software, IEEE, 24(4), 52-60.
7. IC3. (2009). 2009 Internet Crime Report: Internet Crime Complaint Center.
8. ISACA. (2011). Information Systems Audit and Control Association, from http:// www.isaca.org/
9. ISF. (2011). The Information Security Forum - the world's leading independent authority on information security, from https://www.securityforum.org/
10. Kerth, N. L. and Cunningham, W. (1997). Using Patterns to Improve Our Architectural Vision. IEEE Software, 23, 53-59.
11. Kienzle, D. M., Elder, M. C., Tyree, D. and Edwards-Hewitt, J. (2006). Security patterns repository, version 1.0.
12. Kitchenham, B. (2004). Procedures for Perfoming Systematic Review. Joint Technical Report, Software Engineering Group, Department of Computer Scinece Keele University, United Kingdom and Empirical Software Engineering, National ICT Australia Ltd.: Australia.
13. Moral-García, S., Ortiz, R., Moral-Rubio, S., Vela, B., Garzás, J. and Fernández-Medina, E. (2010). A new Pattern Template to Support the Design of Security Architectures. PATTERNS 2010, The second International Conference on Pervasive Patterns and Applications, Lisbon (Portugal). pp. 66-71
14. Okubo, T. and Tanaka, H. (2008). Web security patterns for analysis and design. Proceedings of the 15th Conference on Pattern Languages of Programs, Nashville, Tennessee.

15. Ortiz, R., Moral-García, S., Moral-Rubio, S., Vela, B., Garzás, J. and Fernández-Medina, E. (2010). Applicability of Security Patterns. On the Move to Meaningful Internet Systems: OTM 2010, 6426, 672-684.
16. OSA. (2011). Open Security Architecture, from http://www.opensecurityarchitecture.org/cms/index.php
17. OWASP. (2011). The Open Web Application Security Project, from http://www.owasp.org
18. Rising, L. and Delano, D. E. (1998). The Patterns handbook: Cambridge University Press.
19. Rosado, D. G., Gutiérrez, C., Fernández-Medina, E. and Piattini, M. (2006). Security patterns and requirements for internet-based applications. Internet Research: Electronic Networking Applications and Policy, 16, 519-536.
20. Ryoo, J., Laplante, P. and Kazman, R. (2010). A Methodology for Mining Security Tactics from Security Patterns. HICSS 2010 - the 43rd Hawaii International Conference on System Sciences, Honolulu, Hawaii
21. SANS. (2011). SANS Intitute - Computer Security Training, Network Research & Resources, from http://www.sans.org/
22. Schumacher, M. (2003). Security Engineering with patterns - Origins, Theoretical Model, and New Applications (Vol. LCNS 2754): Springer-Verlag.
23. Schumacher, M., Fernandez-Buglioni, E., Hybertson, D., Buschmann, F. and Sommerlad, P. (2006). Security Patterns: Integrating Security and Systems Engineering: Wiley.
24. Stallings, W. (2007). Network security essentials: applications and standards: Prentice Hall.
25. Yskout, K., Heyman, T., Scandariato, R. and Joosen, W. (2006). An inventory of security patterns. Technical Report CW-469, Katholieke Universiteit Leuven, Department of Computer Science.