Applications of Security Reference Architectures in Distributed Systems: Initial Findings of Systematic Mapping Study

Sajjad Mahmood, Muhammad Jalal Khan and Sajid Anwer
Information and Computer Science Department
King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia
e-mail: [smahmood, g201408880, g201303950]@kfupm.edu.sa

Abstract—There is an increase in use of reference architectures to support software development activities for building distributed systems. Reference architectures are helpful tools to understand and specify functionalizes of a distributed system at a higher abstraction level. From a security standpoint, a distributed system’s reference architecture is one of the potential starting point to study security threats and their characteristics. Both academia and industry have proposed a number of Security Reference Architectures (SRAs), which are reference architectures specifying a conceptual model of security for a system and they provide a mechanism to specify security requirements. The main objective of this work is to investigate and better understand how security reference architecture support building secure distributed software applications. In order to meet our goal, we conducted a systematic mapping study to identify the primary studies related to SRA for distributed software development. We used customized search terms, derived from our research question, to identify literature on SRA for distributed systems. We identified that a significant number of SRAs have been developed first for defense against one or few specific types of security attacks. There is also a focus on developing SRAs to satisfy a security objective during development of distributed systems. Based on the systematic mapping study results, we suggest that there is a need to develop SRAs that help system developers simultaneously enumerate different types of security threats and systematically help to decide where we should add corresponding security patterns to mitigate them.

Keywords—security reference architecture; reference architecture; distributed systems; systematic mapping study.

I. INTRODUCTION

The past several years have seen tremendous changes in distributed software development due to introduction of web 2.0 technologies [14], service oriented architectures [15] and cloud computing systems [16]. These distributed system development technologies have brought with them several new and complex security threats and challenges. To holistically study security of these large and complex distributed systems, we need to start our security analysis from their security reference architectures [1].

A security reference architecture is a reference architecture where security mechanisms have been added in appropriate places to provide some degree of security [1]. Furthermore, a reference architecture is an abstract system architecture that describes system functionalities without any implementation details [1]. Reference architectures are useful to specify main features of a system.

A number of SRAs have been proposed by both academia and industry vendors. For example, Chonka et al. [2] report a technique that is used to observe and discover denial of service attacks against cloud systems. Okuhara et al. [3] report Fujitsu’s security architecture, which logically separates computational environments, authentication and identify management. Similarly, Oracle developed a SRA [4], which addresses data security, fraud detection and compliance with reference to their products.

The literature on SRAs provides a wealth of information on how to analyze security of a system for individual attacks, model system for a security objective(s) or how to help systems meet security compliance requirements of a government organization. For example, Bahmani et al. [5] compared different enterprise information security architecture frameworks with reference to interoperability feature. Lately, Modi et al. [6] presented a survey of intrusion detection techniques in cloud computing systems.

However, there is a lack of systematic investigation of the literature covering SRA in distributed systems. The aim of this systematic mapping study is to collate knowledge to better understand how SRAs have supported system security and identify in what ways it has been applied in the industry.

The remaining of this paper is organized as follows: Section II presents the related work. The research methodology is outlined in Section III. In Section IV, we present and discuss the initial results. Section V discusses the limitation of our study. Finally, the conclusion is presented in Section VI.

II. RELATED WORK

Security is a fundamental concern in any distributed system and a number of security reference architectures have been proposed by industry and researchers’ community. Majority of security reference architectures have been proposed for a particular attack type. There has been significant focus on developing SRAs to mitigate attacks such as denial of service [22], Internet protocol spoofing and denial of service [23].

On the other hand, researchers have also focused on developing security objective specific SRA. For example, Hafner et al. [10] have used enterprise patterns to develop secure services for cloud computing systems. Lombardi and Pietro [11] used virtualization to propose an architecture for cloud protection that monitors middleware integrity.

Even though extensive research has been carried out in the security reference architecture domain, it is necessary to
assess the current state of research and practice, and provide practitioners with evidence that enables fostering future research directions. To the best of our knowledge, there is a lack of systematic investigation of the literature covering SRA in distributed systems.

III. RESEARCH METHODOLOGY

In order to address the research question, we applied Systematic mapping study and literature review [7] approach. A systematic mapping study and literature review is a technique to identify, analyze and interpret relevant published primary studies with reference to a specific research question. Systematic mapping studies are recommended as a review methodology [7] because they allow the researchers to systematically summarize existing evidence from literature, identify research gaps and provide a framework to position future research activities [13].

A systematic mapping study protocol consists of five main phases, as shown in Figure 1. In the first phase of our study, we formulated a research question as follows:

RQ1: How is security reference architecture supporting development of distributed systems?

Next, we constructed the search strategy in line with our research question and performed the search for relevant publications. In the third phase, the identified relevant publications were scrutinized to ensure their relevance. In the fourth phase, the selected studies were evaluated based on the quality assessment criteria. In the last phase, data was extracted from selected studies for further analysis and assessment.

A. Search Strategy

The search strategy for the systematic mapping study is based on the steps as follows:

- Derive the search terms from population, intervention and outcomes.
- Identify alternative spellings and synonyms for major terms.
- Use Boolean ‘OR’ and ‘AND’ operators.
- Verify the derived search term in major academic repositories.

We constructed the following search terms based on our search strategy:

- POPULATION: Distributed systems, Cloud systems, Service-oriented Architecture.
- INTERVENTION: Security Methodology.
- OUTCOME OF RELEVANCE: Different techniques to mitigate security in reference architecture, classification of SRAs.
- EXPERIMENTAL DESIGN: Systematic literature reviews and empirical studies.

We validated our search terms in major academic databases in a scoping study. The following search terms show potential relevance to the research question as follows:

- DISTRIBUTED SYSTEM: Distributed Systems OR Cloud Systems OR Service-oriented Architecture OR Grid Systems; AND
- TECHNIQUE: Technique OR Method OR Model OR Design.

The relevant studies retrieved through the initial search string were used as a guide for the development and validation of the final search string. In the scoping study, we used some relevant publications, which we had previously identified to cross check the validity of the search terms. A broad search was conducted between February 2015 and May 2015 to identify relevant articles published (or available online) up to May 2015.

B. Publication Selection

The following inclusion criteria were used:

- Peer-reviewed studies.
- Papers focus on answering our research question.
- Papers published in English.

We applied the exclusion criteria as follows:

- Papers that are not published in English.
- Papers with no link with the research question.
- Grey publications, that is, papers without bibliographic information.
- In the case of duplicate papers, the most complete version published.

Next, each paper was evaluated against the quality assessment criteria shown in Table 1. Each quality assessment criterion has two answers: ‘Yes’ or ‘No’ with scores of ‘1’ and ‘0’, respectively. The sum of the quality criteria resulted in the quality score for a particular paper. In this study, we only consider publications with a quality score greater than 75%. As a result, 58 papers were finally selected, which met the inclusion and quality assessment criteria.
Initially, when synthesizing the data, data was extracted from the final selection of papers as follows: study details, study research methodology, assessment details and study findings.

IV. RESULTS AND DISCUSSION

The total number of results retrieved using the search terms in the electronic databases are shown in Table 2. After the initial round of screening by reading the title and abstract, seventy one studies belonging to different electronic research databases were selected. After full text readings in the second screening and quality assessment, 58 primary studies were finally selected. Figure 2 shows temporal view of the selected articles from the systematic review, sorted by year of publication. Appendix A presents the primary studies in the review.

<table>
<thead>
<tr>
<th>Quality Criteria</th>
<th>Possible Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is there a rationale for why the study was undertaken? [8]</td>
<td>Yes =1</td>
</tr>
<tr>
<td>Are the research goals clearly stated?</td>
<td>Yes =1</td>
</tr>
<tr>
<td>Is the proposed technique clearly described?</td>
<td>Yes =1</td>
</tr>
<tr>
<td>Is the research empirically validated?</td>
<td>Yes =1</td>
</tr>
<tr>
<td>Are the limitations of this study explicitly discussed? [9]</td>
<td>Yes =1</td>
</tr>
<tr>
<td>Is the study supported by a tool?</td>
<td>Yes =1</td>
</tr>
</tbody>
</table>

Table 2. Search execution

<table>
<thead>
<tr>
<th>Resource</th>
<th>Total Results</th>
<th>Initial Selection</th>
<th>Final Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACM</td>
<td>200</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>IEEE Xplore</td>
<td>427</td>
<td>36</td>
<td>27</td>
</tr>
<tr>
<td>Science Direct</td>
<td>350</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>Springer</td>
<td>61</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>1038</td>
<td>71</td>
<td>58</td>
</tr>
</tbody>
</table>

To answer the research question, the data was carefully extracted and synthesized from the 58 finally selected primary studies. We classified SRAs into four main categories as shown in Table 3.

In our study, the most highly cited category is ‘attack specific SRA’ (60%). Distributed system infrastructure uses virtualization techniques and provide their services through standard internet protocols [6]. Distributed systems are vulnerable to traditional security attacks such as Internet protocol spoofing, routing information protocol attacks, denial of service attacks, etc. Hence, there has been a significant focus on developing SRAs to incorporate specific attack detection and prevention mechanisms in distributed system infrastructure to mitigate security attacks. Table 4 shows a list of popular types of attacks addressed by researchers and industry practitioners.

‘Security objective specific SRA’ is the second highly cited category (reported by 34 % of the articles selected from the systematic mapping study and review). Researchers have also considered developing SRAs oriented to some specific security objectives. For example, Hafner et al. [10] have used enterprise patterns to develop secure services for cloud computing systems. Lombardi and Pietro [11] used virtualization to propose an architecture for cloud protection that monitors middleware integrity. Lately, Fernandez et al. [1] presented a method to build a SRA for cloud systems using security patterns and misuse patterns.

Table 3. List of SRAs Categories

<table>
<thead>
<tr>
<th>Categories</th>
<th>Studies</th>
<th>Count</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry Specific SRA</td>
<td>A21, A26</td>
<td>2</td>
<td>3.44</td>
</tr>
<tr>
<td>Vendor Specific SRA</td>
<td>A16</td>
<td>1</td>
<td>1.72</td>
</tr>
</tbody>
</table>

Furthermore, less frequently cited categories are ‘industry specific SRA’ and ‘vendor specific SRA. There has been couple of SRAs developed for a specific industry. For example, Cohen [12] developed a SRA for industrial control systems. Similarly, Bahmani et al. [5] discussed five enterprise security reference architectures, namely, Gartner
framework [17], SABSA [18], roadmap for information security across the enterprise framework [19], agile governance model based model [20] and intelligent service-oriented enterprise security architecture [21].

It is important to note that over the years, a significant number of SRAs have been developed by industry vendors. All major industrial vendors like IBM, Microsoft, Oracle, Cisco, VMware and Amazon have developed SRAs for their product range. However, in our study, we have not included them as primary studies because most of vendor specific SRAs are available in form of white papers, which do not satisfy our inclusion criteria, as mentioned in Section 2. Hence, we have included only on primary study regarding Fujitsu’s SRA reported by Okuhara et al. [3].

V. LIMITATIONS

Similarly to any systematic mapping study and literature review, our results also depend on the used keywords and the limitations of the search engines. In order to limit the risk of incompleteness in keywords lists, we used alternative spellings and synonyms to build the search terms.

Application of inclusion, exclusion criteria and primary study selection process are also subject to threats to validity of the study. In order to mitigate this threat, all systematic mapping study phases were carried out iteratively with continuous feedback from authors of the paper.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we conducted a systematic mapping study to investigate the use of SRA to support development of distributed systems. Fifty eight studies were finally included, which were further classified into four categories, namely, ‘attack specific SRA’, ‘security objective specific SRA’, ‘industry specific SRA’ and ‘vendor specific SRA’.

Through this systematic mapping study, we identified that researchers have mainly focused on developing SRAs for one or group of individual security attacks. There also has been a focus on developing SRAs oriented to security objectives such as monitoring data and using security patterns to add security mechanisms at appropriate components of a system.

We believe that the results presented in our systematic mapping study and review can be useful for software engineering community as it provides an initial body of knowledge regarding SRAs. In the future, we intend to expand this systematic review to further analyze individual categories and discuss highly cited types of attacks and security objectives in the literature. Another area for future work is to empirically study different industry vendor SRAs and their impact on improving security of distributed systems.

ACKNOWLEDGMENT

The authors would like to thank King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia for continuous support in research. This research is supported by the Deanship of Scientific Research at KFUPM under research grant IN131013.

REFERENCES

APPENDIX A: SYSTEMATIC MAPPING STUDY PRIMARY STUDIES


