ABSTRACT

When considering complex and wide-ranging adaptations of a middleware architecture, there are often several feasible design alternatives. Reasoning in run-time for selecting among such big adaptations would not scale, although implementation of the eventually chosen alternative could be automated. There is therefore a need for early decision support that (in terms of the time needed) can enlighten the effects of the implementation of the various design alternatives. The goal is to integrate the early decision support into the middleware adaptation feedback loop. We propose a decision support method for trade-off analysis considering three aspects: cost, risk, and quality. We have tried out the decision making method in a case study and evaluated the performance of the method with respect to a set of pre-defined success criteria. Apart from the case study itself, the evaluation involved a thought experiment, observations made during the analysis, as well as written feedback. The results of the evaluation indicate practical feasibility of the proposed method. In addition, the evaluation has provided useful insights into strengths and weaknesses of the method and suggested directions for future research and improvements. This paper briefly presents the decision making method, and focuses on reporting the evaluation results.

Categories and Subject Descriptors

D.2.7 [Software Engineering]: Distribution, Maintenance, and Enhancement—Restructuring, reverse engineering, and reengineering; D.2.9 [Software Engineering]: Management—cost estimation, software quality assurance (SQA)

General Terms


Keywords

Adaptation, decision support, impact analysis, risk.

1. INTRODUCTION

Imagine a system architect wanting to adapt a middleware and the system it is a part of, in order to improve the security.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

Analyzing Impacts of Adaptations on Cost, Risk, and Quality
– An Experience Report

Avjot Garcha Singh
University of Oslo
Norway
avjotgs@ifi.uio.no

Aida Omerovic, Franck Chauvel, Nicolas Ferry
SINTEF
Norway
{firstname.lastname}@sintef.no

The architect has, for instance, characterized five possible architecture design alternatives that will provide the required security improvement. All alternatives involve complex and wide-ranging changes on the system. The impacts of the design alternatives, apart from the security improvement, are not trivial to foresee. Although all alternatives are feasible candidates and can be deployed in run-time, there is a need for an early high-level analysis which could help distinguish between the design alternatives. Questions at this stage are: what is the (implementation and operation) cost the design alternatives, what risks are associated with the alternatives, and which effects does each design alternative have on the quality of the overall system? Reasoning in run-time would not provide answers to these questions. For this, we need a light-weight methodological approach that, in terms of the time needed, fits into the middleware adaptation loop. The objective is to facilitate early selection of a preferred design alternative, so that an informed decision is made before deploying the selected one. Thus, we propose an early step in the adaptation loop where the design alternatives are analyzed from a high-level view and ranged based on their estimated impacts on cost, risk and quality. By cost we mean the price to pay for adapting and operating the target system. Risk is defined as "the likelihood of an unwanted incident and its consequence for a specific asset" [17]. According to the International Organization for Standardization [14], quality is defined as "the degree to which a set of inherent characteristics fulfills requirements".

There are already several models for cost estimation (e.g., COCOMO [11] and SLIM [11]). For analysis of security economics, some of the known approaches are Net Present Value (NPV) [10], Return on Security Investment (ROSI) [25], the Cost Benefit Analysis Method (CBAM) [15] and the Security Solution Design Trade-Off Analysis [13]. For risk analysis, some of the well known approaches include CORAS [17], CRAMM [9], OCTAVE [7], Event Tree Analysis [5], Cause-consequence analysis [19] and Fault Tree Analysis [6]. For quality and architecture analysis, approaches such as PREDIQT [20] and ATAM [16] can be applied. The challenge of the existing state of the art is that it either supports analysis of one of the aspects (i.e., cost, risk or quality) at a time in a manner that is practically scalable. A full scale analysis of all three aspects [21] by unifying existing approaches has been proposed. The latter is, however, too complex and time consuming, thus not scaling in a practical setting. There is therefore a need for a practically feasible decision support method which addresses trade-offs between all three aspects: cost, risk and quality.

We propose a light-weight decision making method that fa-
The innovation phase consisted of initial design of the method and requirements related to the method. Thus, the method has, to a certain degree, been developed and evaluated in parallel. In addition to the case study, the evaluation of the method was based on a thought experiment and written feedback from the domain experts – both obtained after the analysis. The objective of the thought experiment was to evaluate the estimates provided by the domain experts. Additionally, observations and notes were made during the case study. Finally, the method was evaluated with respect to a pre-defined set of success criteria.

3. OVERVIEW OF OUR METHOD

In this section, we give an overview of our decision making method, illustrated on Figure 1. The method guides the analyst to systematically model cost, risk, and quality, as well as to correctly use the models.

The process assumes that: the goal of adaptation is characterized; a target system description is made available; and specification of each of the architectural design alternatives is made in advance.

Quality Analysis, Risk Analysis, and Cost Analysis are all independent of each other. In that manner, Quality Analysis, Risk Analysis, and Cost Analysis can be conducted in parallel.

Figure 1: Overview of the decision making method

Input:
- Objective, goal, and vision
- Target system description
- Proposed decision alternatives

Phase 1: Quality Analysis
Step 1: Identification of quality characteristics
Step 2: Development of quality level estimates

Phase 2: Risk Analysis
Step 1: Identification of risk elements
Step 2: Development of risk representations

Phase 3: Cost Analysis
Step 1: Identification of cost elements
Step 2: Development of cost level estimates

Phase 4: Decision Making
Step 1: Specification of selection criteria
Step 2: Graphical visualization of the overall performance
Step 3: Decision making

Output:
- Informed decisions
any order. Quality Analysis involves identification of quality characteristics and estimation of quality characteristic fulfillment associated with the decision alternatives. During the Risk Analysis, the assets and the risks associated with the proposed decision alternatives are identified. Thereafter, consequence and likelihood of each identified risk is estimated. The consequence and the likelihood are specified with respect to a pre-defined scale. During the Cost Analysis, total cost is decomposed into a hierarchical model. The decomposed elements are estimated and aggregated into a total cost representing the expected price of adaptation and operation of the target system.

Our method also includes a final step for trade-off analysis which summarizes the performance of the decision alternatives with respect to cost, risk, and quality. The user assigns selection criteria, which are used for comparing the scores of the decision alternatives. The outcome of this phase is a unified view of the performance of the decision alternatives.

4. OVERVIEW OF THE SETUP AND OUTCOMES OF THE CASE STUDY

The case study was conducted on a system called SensApp [18]. SensApp is an open-source service-based application used to store and exploit data collected by the Internet of Things (IoT). SensApp can register sensors, store the data collected by the sensors and notify registered clients with relevant data and information [18]. The main stakeholders involved in the context of SensApp are: sensor architect, service provider, data miner, and third party application. The sensor architect is responsible for the registration of sensors. The sensors push data to the application, which indirectly triggers the sending of a notification when relevant data is pushed. The service provider is responsible for the operations and maintenance of the sensors. The data miner can query stored data, while a third party application in addition has the ability to register for notifications when relevant data is pushed.

The case study was conducted in an information security context with the objective of improving the security of SensApp. Two analysts and two domain experts participated in the entire case study. The following decision alternatives (DAs) were analyzed:

- DA A: Change in infrastructure
- DA B: Change of topology
- DA C: Change of licenses
- DA D: Change of location
- DA E: Update software

All decision alternatives were specified in detail and considered as realistic alternatives for security improvement. Change in infrastructure means change in the technical architecture needed for the functioning of the service provided by SensApp. Change of topology means change in the configuration of the technical base needed for the functioning of the service provided by SensApp. Change of licenses involves upgrading or purchasing enterprise and commercial software licenses for information security purposes. By change of location we mean geographical relocation of the infrastructure, the platform, and the environment that SensApp is based upon. Updating the current software version of SensApp involves implementation of various security mechanisms in the already existing solution of SensApp.

<table>
<thead>
<tr>
<th>Decision alternative</th>
<th>Total cost (total budget: NOK 1 105 000)</th>
<th>Total quality</th>
<th>Weight of risks</th>
<th>Cost of one point of quality</th>
<th># of catastrophic risk</th>
<th>Risk of one point of quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>720 000</td>
<td>62.5</td>
<td>269</td>
<td>11520.0</td>
<td>2</td>
<td>4.3</td>
</tr>
<tr>
<td>B</td>
<td>740 000</td>
<td>59.7</td>
<td>170</td>
<td>12400.0</td>
<td>0</td>
<td>2.8</td>
</tr>
<tr>
<td>C</td>
<td>810 000</td>
<td>61.1</td>
<td>111</td>
<td>14201.0</td>
<td>0</td>
<td>9.9</td>
</tr>
<tr>
<td>D</td>
<td>890 000</td>
<td>62.7</td>
<td>380</td>
<td>14201.3</td>
<td>0</td>
<td>6.1</td>
</tr>
<tr>
<td>E</td>
<td>855 000</td>
<td>61.8</td>
<td>436</td>
<td>12423,2</td>
<td>0</td>
<td>6.3</td>
</tr>
</tbody>
</table>

Figure 2: Overall performance of the decision alternatives

Table 1 outlines the process undergone during the case study. The first column addresses the specific meeting. The second column specifies the date for the specific meeting. Third column specifies the content for the meeting in question, while the fourth column shows the approximate time in terms of hours spent for the specific meeting, where A denotes the number of hours spent by the analyst and D denotes the number of man hours spent by the domain experts.

The case study was conducted in the period from April to July 2013 and consisted of mainly eight meetings. The first four meetings involved presentation of the target system, i.e., SensApp, and the development of a target system description consisting of UML [8] system models. The target system description ensured that a common understanding of the target system was obtained. The decision making method was conducted thereafter. Note that the last meeting was not part of the analysis itself. The last meeting involved an evaluation of the decision making method based on a thought experiment, followed by a written evaluation after the analysis.

The domain experts had strong technological background and many years of professional experience each. The domain experts were actively participating in the development of the target system description by providing relevant input during these meetings.

As one of the first steps of the target system modeling, the analyst received system documentation of SensApp from the domain experts containing descriptions of the system, architectural models, as well as CloudML [12] and Palladio [22] models of SensApp. The analyst then developed relevant UML models in close collaboration with the domain experts of SensApp by using Enterprise Architect [4]. The first four meetings involved the development of the target system description. The main outcomes of the remaining part of the case study were cost, risk, and quality models of SensApp as is, as well as cost, risk, and quality models with respect to each decision alternative. Finally, the selection criteria were specified and the trade-off analysis performed. A detailed account of all models is available online [1]. Figure 2 shows one of the models, namely the overall performance of the decision alternatives A-E, based on our cost, quality and risk estimates. We have also developed a prototype tool (available online [2]) for modeling the decision alternatives and representing their overall score. The tool is based on Microsoft Office Excel [3] and stores all quantitative input. The tool facilitates the Decision Making phase by defining the selection criteria and enabling propagation of impacts of the input on the overall performance.

5. ASSESSMENT

The assessment of the performance of the method was based on a thought experiment, written feedback and observations. This section summarizes main results from the assessment, while a detailed account of the simulations related to the thought experiment, as well as the full written
feedback, are available online [1].

5.1 Thought Experiment

Four thought experiments were performed in total. The domain experts were provided the specification of SensApp with current cost, risk and quality estimates. For each thought experiment, the analyst presented a specification of the adaptation. Thereafter, effects in the cost, risk and quality estimates were discussed and provided by the domain experts. Independently, from the thought experiment, the analyst performed a simulation of the effects of each adaptation, based on our models from the case study. Eventually, results of the thought experiment and the simulation, were compared.

In the case of the first one, the results gained from the model-based simulation did not correspond with the results gained from the domain experts. In the case of the second one, there was 18% deviation between the estimate obtained from our models, compared to the one provided by the domain experts. In the case of the third and the fourth ones, the results gained through the thought experiment corresponded with the model-based simulation. More details regarding the results of the thought experiment are available online [1].

5.2 Written Feedback

A form for the written feedback was handed out at the end of the last meeting. The main strengths pointed out by the domain experts are the overall simplicity of the method, and the practical feasibility of the decision making method. Both respondents agree upon that the process is quite helpful and meaningful. According to the domain experts, the decision making method was very straightforward and easy to follow.

According to R1: "This decision making process is an interesting endeavor in merging cost/quality trade-off with risk analysis" and "it provides an opportunity to see, side-by-side, numbers quantifying, cost, quality, and risk". "To the best of my knowledge, such an analysis covering risk has never been proposed in the Software Engineering community" (R1).

The main weaknesses pointed out by the respondents are the lack of confidence and uncertainty involved in the estimates provided by the domain experts. Another weak point mentioned by the both respondents is the quantification of distance between the decision alternatives. According to R1: "The two weakest aspects of the process are the definition of time and the completeness of the design alternatives identified in the first place. As far as I understand, time is a key factor here. The time spent implementing a given design alternative obviously impacts the cost, the risk, and the quality of the resulting system. It was not clear to me how we discriminated between a punctual yet very risky design alternative, and, for instance, a long lasting but less risky design alternative".

5.3 Observations Made During the Case Study

In the following we summarize the main observations made by the analyst during the case study.

The target system specification was developed jointly in the group, which ensured that a common understanding was obtained by all participants. The domain experts were actively involved and participated in discussions.

In spite of how detailed the target system description is, the analyst should not develop the models alone. Involving the domain experts during the development of the models ensured that errors and misunderstandings were revealed and resolved. The initial versions of the models contained errors and mistakes, which eventually triggered helpful discussions among the stakeholders involved.

The cost, risk, and quality models were developed in close collaboration with the domain experts. The domain experts had strong technical background, but still the analyst observed that it was necessary to explain the models in order to eliminate confusions.

The estimates provided by the domain experts are highly dependent on the context and circumstances of the target system under analysis. It was therefore necessary to take into account some assumptions during the case study. For instance, the acceptance value for the various quality characteristics would be higher for medical data than for other kinds of data.

Further into the analysis of the decision making method, the domain experts discussed the need of percentages instead of absolute values regarding the acceptance values for the various cost factors. The absolute cost values for software evolution were estimated on the basis of the annual salary of a software developer.

6. DISCUSSION

In this section we discuss, based on the evaluation results, to what extent the success criteria (SC) introduced in Section
2 have been fulfilled. We moreover summarize main threats to validity and reliability.

**SC 1:** The case study, the thought experiment, and the written feedback gained from the domain experts indicate that method facilitates informed decision making in an information security context. The domain experts were able to distinguish between the various design alternatives and their implications with respect to cost, risk, and quality. The written feedback indicated usefulness, but we do not consider our evaluation sufficient for making conclusions regarding usefulness.

**SC 2:** The evaluation indicates that the method can be applied in a realistic setting within limited resources. The results from our case study do indicate practical feasibility of the method. However, SensApp is a small system with limited complexity. It may therefore be argued that SensApp might not have been a realistic target system for our evaluation.

**SC 3:** The stakeholders were actively participating in the analysis and were able to use the decision making method. According to the written feedback from the domain experts, the method was straightforward and fairly easy to understand. However, comprehensibility may vary among the participants depending on the degree of technical background. In this case, the stakeholders involved had a strong technical background. The analyst was also actively involved in the development of the method. Comprehensibility of the method by another analyst would therefore still need to be evaluated. Overall, current results indicate comprehensibility of the decision making method to the stakeholders involved.

**SC 4:** The thought experiment has given some indications of correctness of the models. However, with respect to the evaluation we have conducted, it is difficult to conclude anything about the correctness and certainty of the models. Moreover, the requirements for certainty would vary from case to case (e.g., use of SensApp for handling of medical data would require higher certainty than for another purpose).

**SC 5:** The evaluation indicates that the models were able to capture cost, risk, and quality-related information with respect to SensApp, and support a trade-off analysis. In that manner, the set of models were sufficiently expressive in our specific case. However, on the basis of our evaluation, it is difficult to say if this would be the case in another study.

According to Runeson and Höst [23]: "The validity of a study denotes the trustworthiness of the results, to what extent the results are true and not biased by the researchers' subjective point of view". There was a large number of contextual factors influencing our research, such as: the case study setting, the research method, the target system under analysis, and the participants involved. Our research was highly dependent on subjective judgments provided by the domain experts. As major threat we consider the uncertainty of the models. The domain experts provided estimates based on their experience. In that manner, the validity of the models relies on the confidence to the expert judgments.

The in-depth technological background of the domain experts has to some degree limited the threats to validity. We achieved some triangulation by using multiple sources of evidence in order to strengthen our results. The sources of evidence during the model development have included: target system descriptions, UML system models, presentation slides, meeting notes, written evaluation after the analysis, and observations made during the analysis. During the workshops, the domain experts were actively involved in the model development. This indicates an agreement and common understanding of the method and the models.

A threat to construct validity is the aggregation of cost, risk, and quality, respectively. The aggregation is a rough simplification in our case, and should be addressed in future work.

It is important to emphasize that our case study was based on a small information system with limited complexity. In that manner, it may be argued that SensApp was not a representative case for our research. Thus, the method still needs to be empirically evaluated in order to claim that a repeated case study would give same results (reliability), and that the method would be applicable on other domains (external validity).

It may be argued that there exists a certain correlation between the cost, risk, and quality aspects involved. A specific cost variable might have an internal correlation with another cost variable without the researcher being aware of it. Similar association might also yield for the risk and quality variables involved. Hence, our research lacks the treatment of joint variables, something that may limit internal validity.

Another important fact is that the case study mainly aimed at testing feasibility of the approach. The models developed have not been verified; their main role was 1) to provide an example which demonstrates application of the method and 2) to facilitate further improvement and evaluation of the method. Therefore, the models should not be considered as correct, nor should the case study results be regarded as a risk, quality, or cost analysis of SensApp.

Main challenge in developing of the method was the balancing of the success criteria. Practical usefulness requires that the models are sufficiently informative and correct, at the same time as being easy to understand for a non-expert user. Therefore, we have for the sake of simplicity put some restrictions on the granularity of the models, and the amount of the information being visualized in the last step of the approach. Although our results indicate practical feasibility of the method, further evaluation is needed in order to assess its validity and reliability.

7. CONCLUSION

This work has been motivated by a need for a holistic and light-weight decision support in the context of selection among architecture designs prior to run-time adaptation. When considering several architecture design alternatives, there is a need for better understanding of the impacts in terms of cost, risks, and the system quality. The state of the art offers approaches focusing on one or two of these aspects. Our goal was however to propose a light-weight approach providing a unified view which allows a trade-off analysis of all three aspects at the same time.

We have put forward and tried out a method with human in the loop. The method has been evaluated in a case study targeting a system called SensApp. Apart from the case study, the evaluation involved thought experiment, observations made during the analysis, as well as written and verbal feedback. This paper has mainly focused on reporting the results of the evaluation of feasibility of the method.

The evaluation indicates feasibility in the sense that the method could be applied on a case study and provide useful information regarding the performance of the decision alternatives. We were able to model quality, risk and cost associated with the decision alternatives and conduct a trade-off analysis. The comprehensibility and the expressiveness of the
models seemed to be satisfactory in the context of the case study, while correctness of the models needs further evaluation. The main threat to our findings is validity of the estimates. More evaluation is needed in order to address the threats reliability, but we believe that the method could be useful in the context of adaptation where multiple wide-ranging design alternatives are proposed while only one can be deployed.

Based on our initial experiences, we propose further work on making the method agile and an integrated part of the feedback loop for adaptation. To this end, design alternatives should be detailed enough and the decision models should be sufficiently certain. Thus, future work needs to address uncertainty handling, detailed user guidance, and more empirical evaluation. Future work should also support the sensitivity analysis, i.e., analysis of the degree to which the resulting estimates would change if the input estimates change.

Acknowledgment

This work has been partially supported by the MODAClouds project and the NESSoS network of excellence – both funded by the European Commission within the 7th Framework Programme. It has also been partially supported by the DIAMONDS project funded by the Research Council of Norway.

8. REFERENCES