OPTIMIZING QOS-BASED SERVICE SELECTION IN SERVICE-ORIENTED ARCHITECTURES

Dieter Schuller
Multimedia Communications Lab (KOM), Technische Universität Darmstadt, Rundeturmstr. 10, 64283 Darmstadt, Germany

Keywords: Service-Selection-Problem, Quality of Service, Optimization, Linear Programming, SOA-Security.

Abstract: In Service-oriented Architectures, services can be composed in a loosely coupled manner to realize business processes. Thereby, mentioned services are not necessarily located only within the borders of the own enterprise. In the Internet of Services, multiple service providers offer various services on several service marketplaces. In case services with comparable functionalities but varying quality levels are available at different costs, service requesters can decide, which services from which service providers to select. My research focuses on this service-selection-problem for complex workflows by formulating a linear optimization problem, which can be solved optimally using linear programming techniques. As the actual execution path is probably not known at planning time (e.g., for conditional braches), a worst-case, and an average-case analysis is performed. In addition to considering non-functional, quantitative service properties, I am working in my research towards integrating qualitative service features (as, e.g., security) for different, complex workflow patterns.

1 RESEARCH PROBLEM

In present times, enterprises in highly competitive markets require to execute their business processes and workflows (as IT supported business processes) efficiently. Due to globalization and deregulation of markets, the business processes have to be flexible in addition. This is necessary in order to enable enterprises to react quickly to market driven environmental changes by adapting their business processes. As enterprises often have several legacy systems and a multitude of applications, running on different platforms and operating systems, implemented with various programming languages, the required flexibility is hard to achieve. So, due to this heterogeneity of enterprises' IT architectures, an approach for integrating the applications and legacy systems of the enterprises is necessary to realize the mandatory flexibility and efficiency regarding the enterprises' business processes.

Considering Service-oriented Architectures (SOA), these business processes take center stage. They are realized by composing autonomous, loosely coupled services. Depending on their granularity, they can provide a more or less complex functionality (cf. Krafzig, Banke, Slama, 2004). Thereby, the enterprises do not need to implement all required services on their own. It is possible to rely on the expertise of specialists by invoking services from dedicated, external service providers, enabling enterprises to concentrate on their core business competencies. In fact, cross-organizational business processes gain more and more importance (Leymann, Roller, 2000). In the Internet of Services, various services are offered by different service providers on service marketplaces. In case that functionally equivalent services with varying Quality of Service (QoS) at different cost levels are available, enterprises (as service requesters) can decide, which services from which service providers to select. A certain degree of flexibility and efficiency can be achieved by selecting those (external) services, which meet the enterprise's requirements best with respect to functional and (especially) non-functional service properties. The service selection decision based on QoS properties is addressed in the service-selection-problem.

The service-selection-problem is well known in the literature and discussed by several authors, recently (Yu, Lin, 2005; Anselmi, Ardagna, Cremonesi, 2007; Huang, Lan, Yang, 2009). In this problem, a couple of services (WS – Web service) has to be

Schuller D. (2010).
OPTIMIZING QOS-BASED SERVICE SELECTION IN SERVICE-ORIENTED ARCHITECTURES.
In *Proceedings of the 5th International Conference on Software and Data Technologies*, pages 532-538
DOI: 10.5220/0003047905320538

selected based on their QoS values to realize a given, requested workflow respectively process (PS – Process steps). It is assumed that a set of services, which are able to realize the given workflow with respect to functional requirements, is already identified. This situation is depicted exemplarily for a sequential workflow in Figure 1.

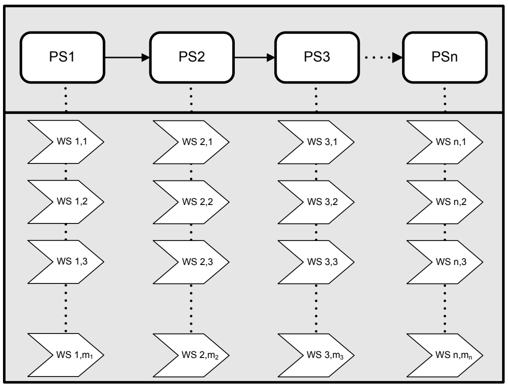

Figure 1: Service-Selection-Problem.

In my research, the service-selection-problem for complex, cross-organizational workflows is focused. For this, a linear optimization problem considering different workflow patterns is formulated, which can be solved optimally using mixed integer linear programming (MILP) techniques from the field of operations research (cf. Domschke, Drexl, 2007), if such a solution exists.

In addition, as the invoked services are probably offered by external service providers, the exchanged (SOAP-) messages (for the communication between the external service provider and the requester) leave the borders of the own enterprise. These messages could be intercepted and manipulated by an attacker (cf. Miede, Nedyalkov, Schuller, Repp, Steinmetz, 2009), potentially leading to failures with respect to the execution of the business processes (cf. Figure 2).

Therefore, security requires to be considered during service selection in order to avoid or at least to recognize service manipulations before the probably modified messages (the results of the service execution) are used and applied in the enterprise's process.

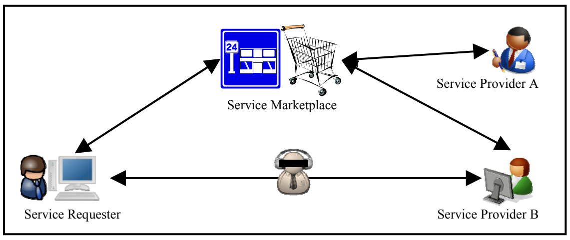

Figure 2: Message Interception.

2 OUTLINE OF OBJECTIVES

My research aims at the development of a smart workflow broker, which accepts (complex) workflow requests in service-based environments, computes an optimal execution plan, performs its execution, and delivers the results back to the initial requester. Thereby, the service selection decision is focused. The question is which services – realizing the requested functionality – from which service providers to select in order to minimize the total costs for the invoked services and to satisfy predefined constraints with respect to QoS (specified by the requester). Further, how can the execution of an optimal execution plan be recovered in case of service failures? Therefore, my research concentrates on the following topics:

- Aggregation functions aggregating the non-functional service properties for (combinations of) complex workflow patterns.
- Linear optimization problem formulation with respect to non-functional service properties for complex workflows, which can be solved by applying MILP, considering a worst-case and an average-case analysis.
- Heuristic algorithms in order to compute almost optimal solutions with acceptable computational efforts, if the number of process steps in the workflow is increasing.
- Quantification of qualitative service properties as, e.g., security in order to be able to consider such properties as quantitative QoS parameters for the optimization.
- Re-planning strategies at design and execution time to (rapidly) react to positive and negative deviations of actual QoS values from the expected ones.

Thereby, a scenario of distributed workflows with a large amount of service providers and service / workflow requesters is considered. My research focuses on the formulation of a linear optimization problem to compute the optimal solution to the service-selection-problem as well as on the development of (efficient) algorithms providing nearly optimal solutions. Here, not only sequential but complex workflow patterns are regarded.

3 EXPECTED OUTCOME

In order to develop the mentioned workflow broker accepting workflow requests and computing – dependent on the workflow length (number of process steps) – (nearly) optimal solutions, the following contributions are planed or already realized.

To start with, an aggregation model for the aggregation of different (aggregation) types of QoS parameters addressing the service-selection-problem for several complex workflow patterns was developed (an extension for further complex workflow patterns is currently ongoing). This model is transformed in a linear optimization problem. Applying MILP techniques, an optimal solution can be computed, if such a solution exists. The respective implementation is currently ongoing. As the considered complex workflow patterns include conditional branches, the ultimately executed path (through the workflow) is not known at planning time. So, it has to be decided, which of the paths to consider for the optimization. If all possible paths require satisfying the predefined constraints, a worst-case analysis is performed. An average-case analysis leads to an execution plan, where the constraints are satisfied in average. In a best-case analysis, only one of the potential paths would have to satisfy the restriction for a certain QoS parameter, which is not of practical interest and will therefore not be considered further.

As the calculation of the optimal solution requires strong computational efforts, heuristic algorithms are currently being developed and implemented as well.

Aside from these contributions regarding the optimization techniques, a first approach for the quantification of qualitative service properties (as, e.g., for security) has been developed. The aim is to be able to consider also qualitative service properties for the service selection by integrating the quantified qualitative values as additional quantitative QoS parameters into the optimization problem. Regarding the projected smart workflow broker, security as qualitative service property is focused. Referring to Figure 3, the Security Module makes sure that (potentially) requested security requirements (e.g., integrity and confidentiality) are satisfied.

To increase the robustness of the broker towards service failures, a re-planning mechanism has been developed and implemented prototypically. Here, a service monitor triggers the interruption of the currently executed service and its replacement – by an alternative (functionally equivalent) service based on semantic annotations – if a negative deviation of the measured form the expected (in Service-Level-Agreements (SLAs) contracted) execution time is detected. This is also done, if problems are encountered with the initial service invocation. The re-planning mechanism – working on execution time as QoS parameter – is planed to be extended for further QoS parameters, for additionally recognizing positive SLA deviations, and for operating supplementary at design time. Here, experienced data about the QoS values of particular services come into play by (rather) considering these values for the optimization instead the ones committed in SLAs.

As already mentioned, the re-planning mechanism is based on semantic annotations. This leads to the presupposition respectively requirement that the service descriptions considered in the smart workflow broker scenario are enriched with semantic information about the services' functionalities. Here, an extension for the prototypically implemented semantic matchmaker is planned.

An architectural blueprint of the targeted smart workflow broker is depicted in Figure 3.

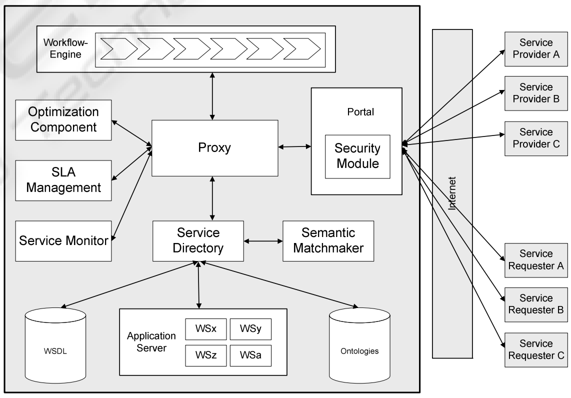

Figure 3: Smart Workflow Broker.

4 METHODOLOGY

For the development of the smart workflow broker, more than one research methodologies are applied. An overview of the contribution and the research methodology used is given in the following list:

- Conceptual work to analyze the research problem and to design the architecture
- Measurements of Web services in a test-bed to record actual QoS behavior for re-planning mechanisms

- Analytical evaluation and optimization of worst-case and average-case service selection as well as analytical models for QoS aggregation
- Simulation to evaluate worst-case and average-case optimization and the re-planning mechanism
- Prototypical implementation of the architecture and the re-planning mechanism

Combining these methodologies allows treating the addressed research questions in an integrated manner.

5 STAGE OF THE RESEARCH

In this section, the contributions of my hitherto research are presented in summary. Here, the service-selection-problem is addressed performing an average-case and a worst-case analysis. In addition, initial concepts have been developed to be able to integrate security and other qualitative QoS parameters into the optimization, and a basic re-planning mechanism has been implemented for sequential workflow patterns.

5.1 QoS Aggregation

In order to solve the service-selection-problem, it is necessary to be able to aggregate the QoS values of the eligible services, i.e. the services capable of realizing the process steps of the considered business process respectively workflow. Otherwise, it is not possible to compare different service selections for the workflow, which is necessary to compute the optimal solution.

How the QoS values have to be aggregated depends on the considered workflow pattern and on the aggregation type of the regarded QoS parameter. E.g., the costs for the whole workflow are calculated by summing up the costs of the selected services, whereas the workflow reliability is determined by multiplying the reliabilities of these services. In my research, the aggregation types summation, multiplication and the Min/Max-operator are considered covering quite all kinds of QoS parameters. In (Schuller, Papageorgiou, Schulte, Eckert, Repp, Steinmetz, 2009), generic aggregation functions for these aggregation types have been developed considering the workflow pattern sequence. In (Schuller, Eckert, Miede, Schulte, Steinmetz, 2010), aggregation functions for the workflow patterns parallel split (AND-split), synchronization (AND-join), exclusive choice (XOR-split), simple merge (XOR-join), and arbitrary cycles (loops) are shown for a worst-case and an average-case analysis.

A specification of further – in my research (regarding the QoS aggregation) yet unconsidered – workflow patterns is available in (Aalst, Hofstede, Kiepuszewski, Barros, 2003).

5.2 Optimization

Based on the abovementioned aggregation functions, a linear optimization problem is formulated in (Schuller et al., 2009) for sequential workflow patterns. In (Schuller et al., 2010), combinations of the workflow patterns parallel split, synchronization, exclusive choice, simple merge, and arbitrary cycles are considered for this purpose. In Figure 4, a (complex) example workflow for a potential combination of these patterns is shown (in Business Process Modeling Notation).

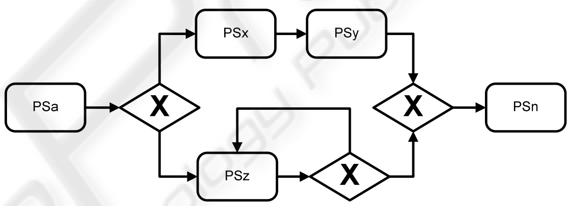

Figure 4: Example Workflow.

The challenge for the formulation of the linear optimization problem lies in the combination of the aggregation functions for the regarded complex workflows and in the linearization of the partly non-linear aggregation functions with respect to the decision variables. As, for instance, the reliability of the eligible services has to be multiplied, the respective decision variables – stating whether the currently considered service is selected or not – are multiplied, too. Having such conditions in the problem formulation, MILP techniques to compute an optimal solution to the service-selection-problem are no longer applicable. Therefore, a linearization step for the mentioned non-linear aggregation functions is mandatory in order to generate linear constraints. For this purpose, an approximation, proposed by (Heckmann, 2008), is applied to the multiplicative non-linear functions. Conditions with the Min/Max-operator are transformed into a set of equivalent linear constraints. As here, the amount of added conditions grows exponentially with the number of process steps within a conditional branch, increasing the restriction strength is proposed in (Schuller et al., 2010).

Having computed the optimal execution plan, aggregating the QoS values of the selected services to calculate the workflow QoS does not lead to exactly one value per QoS parameter, but to a set of values, depending on the executed path. So, due to multiple execution paths, (only) a range of values representing the workflow QoS can be calculated.

5.3 Quantification of Qualitative Service Properties

As the exchanged messages (between service provider and requester) could be intercepted and manipulated by an attacker (cf. Figure 2) if services from external services are invoked, security mechanisms as, e.g., encryption, digital signatures and checksums require to be considered. To integrate these qualitative service features into the optimization, a quantification of these properties is necessary. Therefore, the security features have to be ascertained and consistently assigned to numeric values, to consider them as quantitative QoS parameters for the optimization. By the term "consistently" it is meant that the same value is always assigned to a certain ascertained property. In this context, Jaquith defines a good metric in (Jaquith, 2007) as a "consistent standard for the measurement". It should be possible to express the metric as a numeric value rather than having qualitative labels as "high", "medium" or "low" (Jaquith, 2007).

A first approach to a consistent concept for such a quantification of qualitative QoS parameters – which is at a very early stage – is presented in (Schuller, 2010).

5.4 Re-planning

When invoking one of the services in the execution plan, a service monitor is started, too. As implemented so far, the mentioned monitor measures the time the invoked service takes to provide its functionality. If the measured time exceeds the – in SLAs – contracted execution time (negative SLA deviation), the execution of this service is stopped and a search for alternative services (with respect to functionality) is started. Referring to Figure 3, the service directory is queried to provide alternative services based on semantic annotations. For this purpose, the semantic matchmaker is used. With this set of alternative services, a new execution plan is compiled by solving the service-selection-problem for the remaining (yet unexecuted) process steps. In other words, the initial execution plan is re-planned – a dynamic re-selection of services is performed – taking the already executed services respectively their QoS values into consideration for the adaptation of the restrictions of the (new) optimization problem. The execution of the initial workflow then continues by invoking the (re-planned) services. This re-planning mechanism is presented for sequential workflow patterns in (Schuller et al., 2009).

In case, the described mechanism is repeated several times during workflow execution, applying heuristics is probably advisable due to the mentioned scalability issues when computing an optimal solution.

6 STATE OF THE ART

The main topic of my research, the service-selection-problem, is widely studied in the literature. Heuristic solutions to this problem are proposed and implemented by several authors (Anselmi et al., 2007; Jaeger, Muehl, Golze, 2005; Mabrouk, Georgantas, Issarny, 2009). Tree-based algorithms, where the complex workflow is transformed into a tree-based execution structure (BPTree) based on the alignment of the considered workflow patterns are proposed in (Menascé, Casalicchio, Dubey, 2008). Unfolding arbitrary cycles and computing optimal execution plans for each possible path in the complex workflow is proposed in (Ardagna, Pernici, 2007; Zeng, Benatallah, Ngu, Dumas, Kalagnamam, Chang, 2004). The authors solve the service-selection-problem for complex workflows by formulating and solving the linear optimization problem using MILP techniques solely for (the unfolded) sequential workflows. So, all possible execution paths have to be known and considered consecutively for the computation of an optimal solution.

These shortcomings are overcome in my research, where arbitrary and recursive combinations of (the in section 5.1 mentioned) workflow patterns can be considered for the optimization without a preliminary step for identifying all possible execution paths. The multiple execution paths are directly taken into account when formulating the optimization problem. In other words, an optimal solution to the service-selection-problem is computed considering all execution paths conjunctly. So the execution paths are not considered independently from each other. This is an improvement regarding the mentioned approaches (in the literature) as a consecutive computation for all paths considers the paths independently from each other which probably leads to a case where different services are assigned to the same process step when computing solutions for different execution paths.

7 CONCLUSION AND OUTLOOK

In highly competitive markets, flexible and efficient business process execution is indispensable. Following the paradigm of SOA, using services from (external) service providers (probably having large expertise in a certain business field), enables a flexible and efficient process execution. In case, multiple service providers offer (quite) equivalent services at different quality and cost levels on various service marketplaces (as is probably the case in the Internet of Services), enterprises have the opportunity to decide which services from which service providers to select accomplishing the required functionalities based on their QoS preferences, increasing the achieved efficiency regarding the business process execution.

Therefore, the service-selection-problem for complex workflows is addressed in my research. In this context, I contributed or contribute the following:

- An approach for aggregating QoS values for (combinations of) complex workflow patterns.
- An approach for solving the service-selection-problem optimally (if an optimal solution exists) without preliminary identifying all execution paths.
- A heuristic algorithm for a nearly optimal solution of the service-selection-problem due to scalability issues considering the optimal solution.
- A concept for a quantification of qualitative service properties in order to integrate security into the service selection decision.
- A re-planning strategy to react upon positive and negative SLA deviations for complex workflows.

In particular, the integration of further complex workflow patterns as multi-choice (OR-split) and multi-merge (OR-join) into the optimization problem – yet insufficiently considered (even in the mentioned heuristic approaches) in the literature – and the development of a reasonable concept for quantification of security is currently investigated. The extension of the re-planning mechanism for complex workflows in combination with the progression of the semantic matchmaker (and the therefore required semantic annotations) is addressed as well in my future research.

ACKNOWLEDGEMENTS

This work is supported in part by the E-Finance Lab e. V., Frankfurt am Main, Germany (http://www.efinancelab.com).

REFERENCES

Van der Aalst, W., Ter Hofstede, A., Kiepuszewski, B., Barros, A., 2003. *Workflow Patterns*, Distributed Parallel Databases, 14, 1, pp. 5-51.

Anselmi, J., Ardagna, D., Cremonesi, P., 2007. A QoS-based Selection Approach of Autonomic Grid Services. In *International Conference on Service-oriented Computing*. Vienna, Austria.

Ardagna, D., Pernici, B., 2007. *Adaptive Service Composition in Flexible Processes*, Transactions on Software Engineering, 33, 6, pp. 369-384.

Domschke, W., Drexl, A., 2007. *Einführung in Operations Research*, Springer Verlag. Heidelberg.

Heckmann, O., 2004. *A System-oriented Approach to Efficiency and Quality of Service for Internet Service Providers*, PhD thesis. TU Darmstadt, Fachbereich Informatik.

Huang, A., Lan, C.-W., Yang, S., 2009. *An optimal QoS-based Web service selection scheme,* Information Sciences, 179, pp. 3309-3322.

Jaeger, M., Rojec-Goldmann, G., 2005. SENECA-Simulation of Algorithms for Selection of Web Services for Composition. In *Technologies for E-Services.* Trondheim, Norway.

Jaquith, A., 2007. *Security Metrics: Replacing Fear, Uncertainty, and Doubt*, Addison-Wesley Professional.

Krafzig, D., Banke, K., Slama, D., 2004. *Enterprise SOA: Service-Oriented Architecture Best Practices*, Prentice Hall PTR. New Jersey, USA.

Leymann, F., Roller, D., 2000. *Production Workflow: Concepts and Techniques*, Prentice Hall PTR. New Jersey, USA.

Mabrouk, N., Georgantas, N., Issarny, V., 2009. A Semantic End-to-End QoS Model for Dynamic Service-oriented Environments. In *Principles of Engineering Service Oriented Systems*. Vancouver, Canada.

Menascé, D., Casalicchio, E., Dubey, V., 2008. A Heuristic Approach to optimal Service Selection in Service-oriented Architectures. In *Workshop on Software and Performance*. New Jersey, USA.

Miede, A., Nedyalkov, N., Schuller, D., Repp, N., Steinmetz, R., 2009. Cross-organizational Security – The Service-oriented Difference. In *International Conference on Service-Oriented Computing (ICSOC 2009 Workshops).* Stockholm, Sweden.

Schuller, D., Papageorgiou, A., Schulte, S., Eckert, J., Repp, N., Steinmetz, R., 2009. Process Reliability in Service-oriented Architectures. In *Digital Ecosystems and Technologies*. Istanbul, Turkey.

Schuller, D., 2010. Sichere und Zuverlässige Prozessausführung in Serviceorientierten

Architekturen. In *Central-European Workshop on Services and their Composition*. Berlin, Germany.

Schuller, D., Eckert, J., Miede, A., Schulte, S., Steinmetz, R., 2010. QoS-Aware Service Composition for Complex Workflows. Forthcoming in *International Conference on Internet and Web Applications and Services*. Barcelona, Spain.

Yu, T., Lin, K.-J., 2005. Service Selection Algorithms for Composing Complex Services with Multiple QoS Constraints. In *International Conference on Service-oriented Computing.* Amsterdam, Netherlands.

Zeng, L., Benatallah, B., Ngu, A., Dumas, M., Kalagnanam, J., Chang, H., 2004. *QoS-Aware Middleware for Web Services Composition*, Transactions on Software Engineering, 30, 5, pp. 311-327.