Apostolos Papageorgiou, Jeremias Blendin, André Miede, Julian Eckert, Ralf Steinmetz: *Study and Comparison of Adaptation Mechanisms for Performance Enhancements of Mobile Web Services*. In: Proceedings of the sixth IEEE World Congress on Services (SERVICES 2010), p. 667-670, July 2010, ISBN 978-0-7695-4129-7/10.

Study and Comparison of Adaptation Mechanisms for Performance Enhancements of Mobile Web Service Consumption

Apostolos Papageorgiou, Jeremias Blendin, André Miede, Julian Eckert, Ralf Steinmetz
Multimedia Communications Lab - KOM
Technische Universität Darmstadt
Darmstadt, Germany
apostolos.papageorgiou@kom.tu-darmstadt.de

Abstract — Mobile Web services lie on the intersection of two big IT trends, namely Service-oriented Architectures (SOA) and mobile applications. So, their usage is likely to expand dramatically in the next years. However, the heavyweight nature of service-orientation in terms of the messaging-overhead that is necessary in order to achieve interoperability and loose-coupling comes in contrast with the lightweight nature of mobile devices and with the need to transmit wirelessly as few data as possible. This study categorizes the mechanisms that have been designed in order to bridge this gap, provides comparisons, discusses the results of related experiments, and introduces the future scenario in which the insights of the study can be exploited.

Keywords – Mobile Web services, Mobile SOA, Performance, QoS, Adaptation

I. Introduction and Motivation

Web services, so far the most common technology for realizing the SOA paradigm, start appearing in global marketplaces, leading to the so-called "Internet of Services" (IoS), while *mobile Web services* play an important role among them. With this term we refer to Web services which are consumed from mobile devices, in addition to other potential clients. A survey of TechTarget [6] indicated the predominant types of service-based applications planned for the future, positioning "Web services for mobile apps" at the second place of the related list with 60%, even higher than "composite application assembly" (58%), which has been the "flag"-potential of SOA. For these reasons, research efforts have already focused on performance enhancements for mobile Web services.

In the context of a single service-oriented application, the usage pattern of services may be well defined, so that any performance optimizations can be foreseen and handled at design-time. In the context of the IoS, however, many services will be developed and published without knowing how their usage will look like, or what kind of clients will dominate their consumption. So, performance adaptations will be desired during their operation.

In the case of mobility, there are some techniques which may produce pure overhead when applied to simple service consumption, but which significantly enhance Quality of Service (QoS) under certain circumstances. These circumstances are related to network characteristics, client features, and more. Before we analyze these techniques in later sections, we explain the main *reasons why service consumption is subject to different rules when the consumer is a mobile device*.

- Wireless connection features, such as *limited bandwidth and intermittent connectivity*, can have the following effect: approaches for the service implementation or consumption that perform better or equally to other approaches in a fixed network, can perform worse in a particular wireless network. For example, experiments of [16] proved this for the case of compression. So, a SOA platform may not be able to know at design-time which is the most suitable approach for offering a service.
- Device constraints, like *limited CPU or memory*, have similar consequences, but while connection constraints may favour approaches that transmit less data, device constraints favour approaches that need less processing. This leads to tradeoffs and complicates decisions. Despite technological progress, the gap will continue to exist and the mentioned differences will remain valid for two reasons: Firstly, because new constrained devices, e.g. sensor nodes, will become capable of consuming Web Services, and, secondly, because the workload and the overhead of new services grows in parallel to device improvements. Another explanation why this gap will not be bridged is given in [4].
- Concerns about *energy*, as well as *radiation*, have initiated a very big number of IT projects with environmental targets, e.g., aiming to minimize energy consumption or radio-activity, possibly trading-off others aspects, such as costs or performance.

With this regard, Section 2 refers to our previous work in order to explain the role, the position, and the nature of the adaptation mechanisms for mobile Web Services inside the IoS. Section 3 states the purpose and the contributions of our study. Section 4 analyzes and categorizes the state-of-the-art mechanisms from which adaptation mechanisms for mobile Web Services can be derived, while Section 5 explains what kind of experiments and comparisons accompanied this study. A summary and the future work which can build on the results of this study are given in Section 6.

II. OUR WEB SERVICE ADAPTATION SCENARIO

Our concerns about estimating the suitability of different techniques for mobile Web Service adaptation stem from the needs of our "service marketplace mobility-extension", called Mobility Mediation Layer (MML). The MML is a framework for bridging the gaps towards mobile SOA. In [12], a study of the challenges and the potentials of mobile services indicated the need for the MML as a service marketplace extension that performs mobility-oriented tasks, such as the compatibility-support, the automatic context-enrichment, and more. Among others, the MML facilitates the re-offering of existing services with more lightweight interfaces for performance reasons. As coarsely shown in Fig.1, proxies can be generated for services that need to be offered wirelessly in a lightweight manner. They could perform protocol-transformation, compression or any other of the techniques that will be listed in Section 4.

After the enforcement of the adaptation mechanisms, the mobile developers will have further possibilities for programming clients that use the "adapted" service, in addition to the possibility of directly calling the Web Service. The generated interfaces will be offered and published in the marketplace by the MML itself, which can also be seen as a business model for the MML.

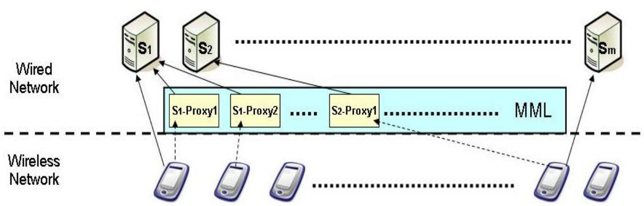

Figure 1. Adapting mobile Web Services in the MML.

III. RELATED WORK AND CONTRIBUTION OF OUR STUDY

Both the design-time decision about the set of alternative access methods that the MML should be able to offer and the runtime decision about which of these alternatives should be put into action for which service, have to be based on a detailed knowledge of measured benefits, which appear in any possible case. To our best knowledge, no study or survey has gathered and directly compared the adaptation mechanisms for mobile Web services so far, and, indeed, no existing work proved suitable to form the basis for our future plans.

Concerning the adaptation mechanisms themselves, the related work is completely covered in the next section. As for related studies, we mention the work of Lonthoff and Ortner [10], which lists and describes some generic approaches. Nevertheless, it is limited to descriptions and does not offer deeper analysis or comparisons. So, *main contribution* of the work at hand is the analysis of adaptation mechanisms according to the circumstances that make them beneficial and the mapping of related work to identified categories. After a detailed study of related work, we consider the following points as further contributions:

- Two-fold purpose experiments: validation of some results found in literature, as well as measurements under a technical setting that had not been examined yet, namely the variation of the used data types.
- Description of the Web service adaptation in the MML as a future scenario where adaptation mechanisms for mobile Web Services can be hosted.

IV. ANALYSIS AND CATEGORIZATION

The consumption of Web services by mobile devices has many aspects that affect QoS. After a careful literature analysis among the approaches that aim at reducing the overhead of mobile Web service consumption, we identified four components that comprise the technical setting of a mobile Web service usage scenario. These are the service itself, the mobile device, the network connection and the application that uses the service. Each of them has various aspects that determine which approach would perform better. Table 1 lists the components, the aspects, and their possible values. The possible values were chosen with the minimum granularity that is necessary for our analysis. For most aspects, the values *small* (s), *medium* (m) and *high* (h) (or *large* (l)) were a good fit. For others, the possible values are *yes* (y) and *no* (n), instead. Example values (thresholds) are given for some aspects. These thresholds are not critical for the analysis and may depend on the goals of the developer or the platform that uses this knowledge. They are provided here just in order to give a feeling about the meaning of s, m, and h (or l) in the context of modern systems.

Our study among the approaches that try to optimize mobile Web service communication showed that most of them aim at improving response times, usually by minimizing the amount of transferred data. A number of these approaches are listed in Table I. Each one is described by a usage setting in which its authors/developers expect to have a big benefit. We are limited to this performance comparison. Further aspects like possible client incompatibilities (e.g., to RMI), decision algorithms for the proxy generation, legal issues of re-providing services etc., concern the logic and the constraints of the decision process, and rather remain subject of future work. An example of how the approaches were analyzed follows, while aspects that proved to be irrelevant to the efficiency were each time dropped from the table.

In [9] and [13], for example, "SOAP-over-UDP" (SoU) is compared to "SOAP-over-HTTP" and "SOAP-over-TCP". The transmission overhead of the three methods is compared, showing that SoU has the smallest. Still, the improvement is highest for small data volumes, while it gets very small for high data volumes. So, the biggest improvement is achieved with small and medium data volumes. Also the execution time is compared through a loopback link (server and client on the same machine) and over a WLAN link. On the loopback link, the SoU improvement is small, whereas it is ten times bigger over the WLAN link, which introduced a big latency. However, the UDP protocol does not guarantee a failure-free data transmission; hence the application should be able to cope with this feature.

TABLE I. ASPECTS THAT COMPRISE THE TECHNICAL SETTING DURING THE USAGE OF A MOBILE WEB SERVICE

Component	Aspect	Values
Service	User data size to SOAP message size ratio	small, medium, large
	SOAP message size	small: ≤100b, medium: ≤1kb, large: >1kb
	Caching possible	No, Session, User, Everyone
	Statefulness	Yes, no
	Processing time	small: ≤1s, medium: ≤1minute, high: >1minute
	Connection-Setup direction	Consumer→Provider, Provider→Consumer
Application	Service call frequency	small: ≤1 per 10minutes, medium: ≤100 per 10minutes, high: >100 per 10 minutes
	Service call dependence	small, medium, high
Device	CPU Power	small: mobile sensor, medium: smartphone, large: laptop
	Available Memory	small, medium, large
Device Network Connection	Bandwidth	small: ≤50kb/s, medium: ≤1mb/s, high: >1mb/s
	Latency	small: ≤10ms, medium: ≤50ms, high: >50ms
	Packet loss	None: ≤0.01%,small: ≤5%, medium: ≤15%, high: >15%
	Stability	small, medium, high
	Disconnected periods	Yes, no
	Directionality of connection establishment	Device→Internet, no

TABLE II. ANALYSIS OF APPROACHES FOR ADAPTED WEB SERVICE CONSUMPTION ACCORDING TO THE SETTINGS THAT MAKE THEM BENEFICIAL

	Component	Device Network Connection					Device	Service			Application		
	Approach \ Aspect	Bandwidth	Latency	Packet loss	Stability	Disconnected peroids	CPU Power	Data size / SOAP size	SOAP message size	Processing time	Service call frequency	Service call dependence	Expected response time Improvement over SOAP/HTTP/TCP
[9],[13]	SOAP-over-UDP		≥m	≤s					≥m	≤s		≤s	8-10x
[17]	SOAP-over-SCTP			≥m	≥h						≥h		1.1x-1.3x
[8],[16]	zLib Compression	≤s					≥m		≥m				1x-1.5x
[7]	SOAP-over-WAP	≤m											1.3x
[8]	SOAP-over-TCP with persistent connection	≥m			≥h				≥m		≥h		2x-5x
[3]	SOAP-over-SMTP/POP3				≤m	y				≥h			-
[2]	Wireless SOAP	≤m						≤s	≥m				3x-5x
[5]	JAVA RMI	≥m	≥m	≤s									>10x
[11]	HHFR	≤m						≤m			≥m		1.5x-10x
[1]	MundoCore RMC						≤m						3x-5x
[14]	Fast Web Services	≤m						≤m	≥m				2x-10x

A similar logic, easily traceable through a careful analysis of Table II, was followed for the remaining approaches. In addition, *the performance enhancement expected by the authors/developers of the approach* is given in the last column. This is expressed in comparison to standard SOAP calls and varies a lot even for the same approach, as it is exactly this value that depends on the other aspects. Other approaches exist, e.g. REST-based. However, some of them were excluded from the results because of lack of reliable and well-documented published material.

At first sight, no strict grouping seems to be possible based only on the values of the aspects, so the presented analysis remains our main contribution. But instead of categories that are strictly based on “beneficial circumstances”, we list five “weakly” defined groups, which are obtained based on the ideas behind the approaches. Up to an extent, approaches of the same group are expected to offer benefits under similar circumstances. The following distinguishable ideas were found in the studied approaches:

- Reduction of redundancies in a stream of Web service calls ([8], [11], [17]).
- Replacement of SOAP calls with alternative RPC Mechanisms ([1], [5]).
- Reduction of the amount of data transmitted during a Web service call, while leaving the rest of the Web service stack untouched ([2], [8], [7], [14], [16]).
- Introduction of a message queuing infrastructure that queues SOAP messages for retrieval when the device connection is ready for it [3].
- Ignorance of transmission failures, probably for non-critical service calls ([9], [13]).

V. CURRENT EXPERIMENTS AND OUTLOOK

A couple of the examined approaches have been experimentally validated and further examined. Experiments were done with implementations that correspond to the approaches of [5] and [16]. We measured the data overhead reduction and the response times over simulated wireless networks. Most important and novel, however, is the variation of an aspect that has not been examined in previous experiments, namely the data types used by the service.

In Fig. 2 we provide the results of the experiments that measured the relative overhead reduction of a proxy-based approach that uses protocol-translation as in [5] (RMI) for the wireless transfer. The achieved proportional data volume reduction (size of transferred data / size of SOAP data) was measured both for a service that uses single types and for a service that uses complex types. In both cases we varied the volume of the actual data. The experiment has been repeated many times, though repeated tests are necessary only when response times are measured as well.

As the results prove, the protocol-transformation offers no benefit for services that use big-sized single data (e.g., images) and a bigger, varying benefit in other cases. Similar experiments show that compression-based approaches lead to different curves, offering benefit in cases where protocol-translation does not, and vice versa. Further experiments, which focused on different varying features, e.g., network connections, as well as on different goals, e.g., response time reduction, are not presented due to space constraints.

Our future work is driven by two main goals. The first goal is to choose a set of approaches that, according to our study, covers as many different technical settings as possible, or the most common of them, and implement a Web Service Proxy Generator that can automatically offer the interfaces that correspond to them. The second goal is the further exploitation of the insights of this study in order to investigate decision algorithms that specify which approach/interface should be chosen each time. These decisions could depend not only on the circumstances and the technical settings but also on any other information of the service marketplace that could assist this decision, e.g., past user decisions ("Quality-of-Experience") and market trends.

ACKNOWLEDGMENT

This work was funded in part by means of the German Federal Ministry of Economy and Technology under the promotional reference "01MQ09016" ("Green Mobility" Project). The authors take the responsibility for the contents.

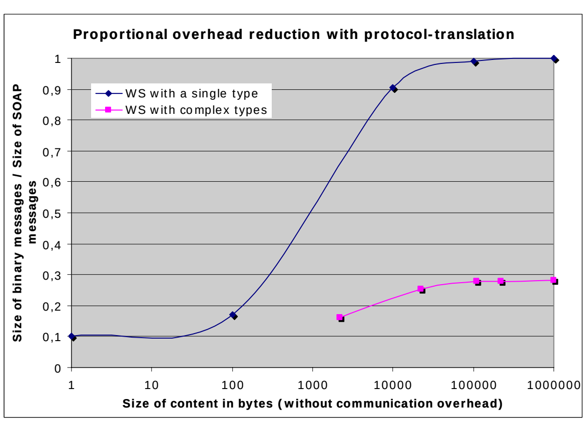

Figure 2. Experimental result showing how the adaptation benefit may depend on the data types used by the service

REFERENCES

[1] E. Aitenbichler, J. Kangasharju, and M. Mühlhäuser, "Mundocore: A lightweight infrastructure for pervasive computing," Pervasive and Mobile Computing, vol. 3, August 2007, pp. 332-361, doi: 10.1016/j.pmcj.2007.04.002.

[2] N. Apte, K. Deutsch, and R Jain, "Wireless SOAP: optimizations for mobile wireless web services," Proc. International World Wide Web Conference (WWW '05), ACM Press, 2005, pp. 1178-1179, doi: 10.1145/1062745.1062927.

[3] A. Buschmann, C. Werner, S. Schmidt, and S. Fischer, „Protokollunterstützung für SOAP Web Services auf Mobilen Geräten," Praxis der Informationsverarbeitung und Kommunikation (PIK), vol. 27, 2004, pp. 140-144, doi:10.1515/PIKO.2004.140.

[4] C. Canali, M. Colajanni, R. Lancellotti, "Performance Evolution of Mobile Web-based Services", IEEE Internet Computing, vol. 13, March 2009, pp. 60-68, doi:10.1109/MIC.2009.43.

[5] D. Davis and M. Parashar, "Latency Performance of SOAP Implementations," Proc. 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID '02), IEEE Computer Society, 2002, pp. 407-412, ISBN:0-7695-1582-7.

[6] C. Frye, "SOA Growth and Change", TechTarget survey, available at http://searchsoa.techtarget.com/news/article/0,289142,sid26_gci1351532,00.html, last updated on March 2009, last visited on April 2010.

[7] G. Gehlen and R. Bergs, "Performance of mobile web service access using the wireless application protocol (WAP)," Proc. World Wireless Congress, 2004, pp. 427-432, ISSN:1529-2592.

[8] J. Kangasharju, S. Tarkoma, and K. Raatikainen, "Comparing SOAP Performance for Various Encodings, Protocols, and Connections," Lecture Notes in Computer Science (LNCS), vol. 2775, September 2003, pp. 397-406, doi:10.1007/b12004.

[9] K. Lai, T. Phan, and Z. Tari, "Efficient SOAP Binding for Mobile Web Services," Proc. IEEE Conference On Local Computer Networks (LCN '05), IEEE Computer Society, 2005, pp. 218-225, doi:10.1109/LCN.2005.62.

[10] J. Lonthoff and E. Ortner, "Klassifikations- und Lösungsansätze für Web Services im mobilen Umfeld," Proc. 2. Konferenz für Mobilität und Mobile Informationssysteme (MMS '07), March 2007, pp 73-84.

[11] S. Oh and G. C. Fox, "Optimizing Web Service Messaging Performance in Mobile Computing," Future Generation Computer Systems, vol. 23, May 2007, pp. 623-632, doi:10.1016/j.future.2006.10.004.

[12] A. Papageorgiou, B. Leferink, J. Eckert, N. Repp, and R. Steinmetz, "Bridging the Gaps towards Structured Mobile SOA," Proc. 7th International Conference on Advances in Mobile Computing & Multimedia (MoMM '09), ACM Press, December 2009, pp. 288-294, ISBN: 978-1-60558-659-5

[13] K. A. Phan, Z. Tari, and P. Bertok, "A Benchmark on SOAP's Transport Protocols Performance for Mobile Applications," Proc. ACM Symposium on Applied Computing (SAC '06), ACM Press, 2006, pp. 1139-1144, doi:10.1145/1141277.1141548.

[14] P. Sandoz et al, "Fast Web Services". Sun Developer Network, http://java.sun.com/developer/technicalArticles/WebServices/fastWS/, last updated on August 2003, last visited on April 2010.

[15] J. Shi, Y. Jin, H. Huang, and D. Zhang, "Experimental Performance Studies of SCTP in Wireless Access Networks," Proc. IEEE International Conference on Communication Technology (ICCT '03), IEEE Press, January 2003, pp. 392-395, doi:10.1109/ICCT.2003.1209105.

[16] M. Tian, T. Voigt, T. Naumowicz, H. Ritter, and J. Schiller, "Performance Considerations for Mobile Web Services," Computer Communications, vol. 27, July 2004, pp. 1097-1105, doi:10.1016/j.comcom.2004.01.015.

[17] N. Wang, M. Welzl, and L. Zhang, "A High Performance SOAP Engine for Grid Computing," ACM Networks for Grid Applications (GridNets), vol. 2, 2009, pp. 1-8, doi:10.1007/978-3-642-02080-3_1.