Towards QoS-awareness of Context-aware Mobile Applications and Services

Katarzyna Wac

Abstract. In our current connected wireless world, mobile devices are enabled to use various networking facilities. Although this enables mobile users to communicate *any time and any place*, it may also be very intrusive. There is a high need to manage the information stream a user receives on his/her mobile device. Context-awareness seems to be a promising way to manage this information stream and to provide the means to communicate at *the right time in the right way*.
Current context-aware applications benefit from the user context (e.g. location information), however, they do not consider the quality of service (QoS) offered by various networks (i.e. only best-effort QoS is considered). The research discussed in this paper focuses on a QoS- and context-aware service infrastructure supporting the development of mobile applications in a heterogeneous network environment. We argue that the use of context information helps to better capture the user's required QoS and improves the delivered QoS.

1 Introduction

The emergence of new wireless broadband networks and diverse miniaturized and personalized networked devices, give rise to variety of new mobile services in our daily life. Ultimately, these mobile services are executed as we move: in different places, at different time and under different conditions. Hence, these services get a continuously changing information flow from their execution environment. The management of this flow becomes vital for mobile services delivery. This means that a communication paradigm needs to shift from *any time and any place* into *the right time in the right way*, as the former may be very intrusive. Context-awareness is a promising way to manage the information flow, as context is any information that characterizes user's environment and situation, (e.g. location, time), and any object relevant to the interaction between the user and a mobile service [1].

For any mobile service the underlying communication is provided by heterogeneous network environment; consisting of wireless and wired networks. Each network is responsible for a section of the 'end-to-end' communication path between a mobile user and application server placed in a service provider network (Figure 1[1]).

[1] The mobile operator network comprises wireless access network and wired core network.

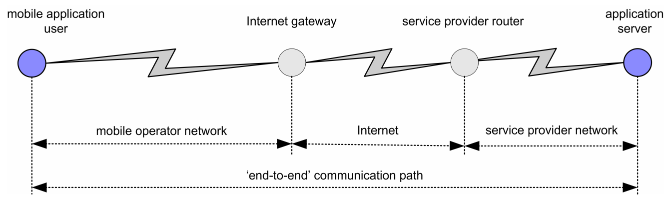

Figure 1. An 'end-to-end' communication path for a mobile service delivery

Mobile connectivity, i.e., persistency of a wireless connection during the act of being mobile, and *quality of service* (QoS) *offered* by this connection, are critical factors for success of mobile service delivery. However, current rules for a connectivity choice are rather simple - a default wireless connection is chosen at a service-design time and assumptions are made regarding its offered QoS. Because a wireless link is usually a bottleneck in the end-to-end communication path, the assumptions regarding its offered QoS imply the assumption for the end-to-end offered QoS. Consequently, current mobile services are *delivered* with a best-effort (end-to-end) *quality*, and without consideration of the mobile user's *required QoS*.

The complication rises from the fact that nowadays various connectivity possibilities coexist offering different QoS. However, current mobile applications are unaware of that - the wireless network is chosen based on its availability, and not its offered QoS. Moreover, the mobile user's QoS requirements are not really considered or assumed to be static and attempted to be met with a best-effort service. Another complication is that the mobile application does not 'learn' from the end-to-end QoS experienced along the service delivery. This QoS information is not logged to be used further for QoS predictions, neither for a user himself nor for the other mobile users.

With respect to all given complications, we research on a QoS- and context-aware service infrastructure that supports the development of mobile applications in a heterogeneous network environment. We argue that mobile service infrastructure must not only be context-aware, but also QoS-aware; aware of the user's required QoS and QoS offered by various networks at the user's location and time. A wireless connection must be chosen with respect to what is required and what is offered. If necessary, service delivery should be adapted to what is offered (e.g. by means of changing the application protocol parameters). We argue that QoS- and context-awareness improves delivered QoS and that context-awareness helps to better capture the user context-depended QoS requirements.

Furthermore, we take the concept of QoS-awareness even broader and indicate that the actually delivered end-to-end QoS must be continuously logged by the mobile service infrastructure, and further used for QoS predictions. That leads to the proactive QoS-aware and context-aware infrastructure. We point necessity of development of a *QoS-context source*. It carries the responsibility of accumulation of logs on delivered end-to-end QoS from different mobile service users and provision of predictions of this QoS (i.e. along the particular trajectory traversed in a particular timeframe) to mobile applications. Hence, we aim in development of a 'route navigator' (i.e., outdoor and indoor GIS-based system) enriched with a QoS prediction for a particular user's trajectory.

The following section gives an overview of the existing research related to the topic. Section 3 provides an explanation of our research context and section 4 -

research trajectory within this context. The state of our thesis is discussed in section 5. The conclusion in section 6 summarizes our thesis goals and provides a critical evaluation of the work.

2 State-of-the-art

There has already been a lot of research on QoS and context-awareness as separate topics, and this section presents only work the most important for us.

Projects like Equanet [3] developed a modeling-based performance evaluation method of the end-to-end QoS delivered to a mobile user over heterogeneous communication networks. In contrary, we consider measurements-based evaluation methods. Moreover, this project only focuses on VoIP and mobile web browsing mobile services, while we do not put constraints on the type of considered services. The CELLO project [4] concentrates on the location-based performance assessment of wireless communication infrastructures. Data regarding networks' performance is stored in a GIS system. However, as a performance indicator this project only considers signal strength and not the end-to-end QoS, as we do. Moreover, the overall goal of the project is to enhance the mobile operator network; the data is not used for mobile users, as we propose in our research. The publication of [5] provides a framework for network-aware applications and indicates that an application-level monitoring is one of the methods for application to be network-aware. However, this publication only considers the end-to-end bandwidth, and no other QoS parameters, as we propose. Similarly, [6] and [7] provide an idea on network resource awareness at the application level. Both indicate user context as necessary information for an application to adapt. However, both indicate the wireless access network and not end-to-end resources availability, as we do. Moreover, the mobile connectivity context source indicated in [7] is based on the single user's history of connectivity, and is used for a user himself to derive further context information. Hence, it will not be used for other users, as we indicate in our research.

Based on this short representation of the ongoing research, we define the innovative contribution of our thesis to the existing state of the art as the fact that we take the end-to-end QoS characteristics as context information and we introduce QoS-context source created based on information acquired from users and used for users.

3 Research context

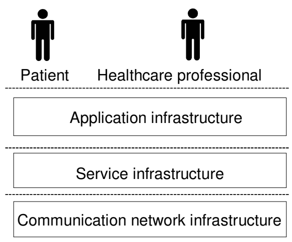

Figure 2. AWARENESS architecture

The context of our thesis is provided by the Dutch national Freeband AWARENESS project [8]. This project research an infrastructure that supports the development of context-aware and pro-active services and applications and validates it through prototyping in mobile healthcare (i.e., m-health) domain.

AWARENESS defines a three-layered architecture. The bottom layer is the *communication network infrastructure* layer, offering seamless mobile connectivity (e.g. 2.5G/3G/WLAN). This layer spans the 'end-to-end' communication path indicated in Figure 1. The middle layer is the *service infrastructure* layer providing an execution environment for mobile services (e.g. service discovery and service context management functions). The top layer is the *application* layer offering *generic application services* like application-level context management and *domain specific services* like health tele-monitoring services, e.g., for the epilepsy, spasticity and chronic pain domains.

We position our research vertically across these three layers. We will incorporate the QoS requirements of service users and map them into system QoS requirements – into the requirements at the service infrastructure layer and then into the requirements at communication network infrastructure layer. To merge QoS-awareness and context-awareness, the context management function at the service infrastructure layer will be enriched with the management of QoS-context information. Moreover, we will develop a QoS-context source interacting with the service infrastructure layer.[2]

4 Research trajectory

Our research trajectory consists of a few consecutive phases: 1) analysis of research-related concepts (to identify the current situation and its problems) as a basis for formulation of our research questions, 2) putting forward a hypothesis on a possible solution and 3) defending the hypothesis, eventually proving it to be a valid theory. Figure 3 presents our research trajectory cycle and particular actions taken in each of the phases. All phases may be repeated cyclically, if necessary.

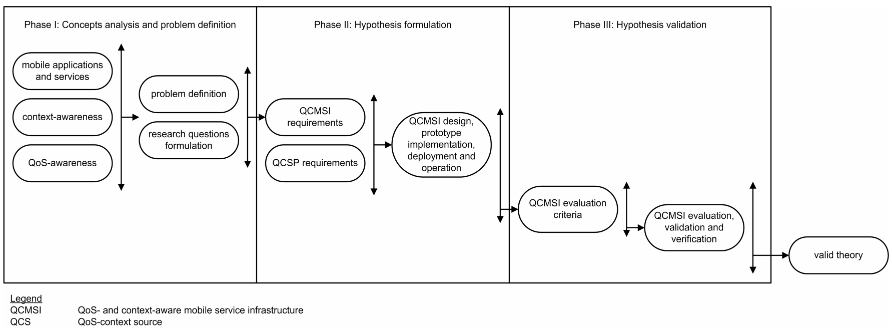

Figure 3. Research trajectory

Phase I In Phase I we start from the system related-concepts analysis and problem definition. We consider three main areas of literature review relevant for this research: a) mobile applications and services (e.g., current m-health applications), b) context-

[2] I gratefully thank Dr Aart van Halteren for his supervision. This is an ongoing research in frame of the Dutch Freeband AWARENESS project (BSIK 03025), partially supported by the E-NEXT Network of Excellence (FP6 IST 506869).

awareness (e.g., user/system context, current context-aware services and toolkits) and c) aspects of the QoS-awareness (e.g., what is QoS for mobile services, QoS management). This study results in a problem analysis of current context-aware service infrastructures and their QoS-(un)awareness.

Furthermore, Phase I aims in deriving specific research questions based on understanding the existing service infrastructure and its problems.

Phase II Following the research questions, in Phase II we will derive a hypothesis of a possible solution for the identified problem. As we stated earlier, we research on a QoS- and context-aware service infrastructure that supports the development of mobile applications in a heterogeneous network environment. Due to the nature of the problem, we derive our hypothesis based on the repeatable service lifecycle model [9], consisting of the following phases: a) requirements (i.e., set of features service must conform; service user and system[3] requirements), b) architectural design (describing the service organization in terms of structural elements, their composition, their interfaces and behavior), c) implementation (hardware, software and firmware used), d) deployment (service availability to users) and e) service operational phase (service maintenance).

Following this model, we will analyze requirements for: the QoS- and context-aware service infrastructure and for a QoS-context source. The former dictates the requirements for the latter. Based on these requirements, we will propose an architectural design for the QoS- and context-aware service infrastructure and its interface with QoS-context source. Following the AWARENESS project goals, prototype implementation and deployment of the system, will be done in m-health domain.

Phase III The system prototype will aim in proving such whether the proposed solution fits the identified earlier problem (Phase I) and conforms the identified requirements (Phase II). Hence in Phase III we will attempt to defend (or refute!) our hypothesis. Firstly, we will derive evaluation criteria along which the prototyped system will be validated and verified. The evaluation will be executed with real m-health service users and will aim to prove that use of context information indeed helps better capture the user's required QoS and improves the delivered QoS. If we defend our hypothesis to be a theory, then we would like to indicate the utility of our solution for a mobile service in any application domain (e.g., commerce, entertainment).

5 Current work

The research discussed in this paper started in November 2004 and consequently is in a starting phase. Some activities from Phase I are presented in the first sections of this paper and continued in this section together with the initial ideas for Phase II.

5.1 Research questions

We research on context-aware mobile service infrastructure to enrich it with QoS-awareness. Moreover, we research on a QoS-context source. Therefore, we define the following research questions:

1) What are mobile user's end-to-end QoS requirements? How to translate them into

[3] A system delivers a service; a service is an external observable behavior of the system.

the system requirements?
2) What end-to-end QoS context information is required at the context-aware service infrastructure level, how to get, and how to use it? How context may improve end-to-end QoS actually delivered to a mobile user?
3) What are the requirements for, and how to develop a QoS-context source created by users for users?
4) Based on the data aggregated in QoS-context source, what algorithms are used to predict end-to-end QoS along the mobile user's trajectory traversed in a given timeframe?

5.2 Epilepsy tele-monitoring scenario

The following future application scenario illustrates how an epileptic patient can benefit from context- and QoS-aware mobile health care service. The scenario (in boxed paragraphs), follow explanations of the technology we propose to support it.

Sophie (28) is since four years an epileptic patient living in the Paris suburbs. Epilepsy is a neurological disorder, in which brain nerve cells release abnormal electrical impulses so-called seizures. Although the occurrence of a seizure is sudden and unexpected, she does not feel limited in her active life because she is treated under a continuous healthcare program; the Epilepsy Safety System (ESS) tele-monitors her health.

The ESS (Figure 4) is a distributed system, responsible for predicting, detecting and handling the occurrence of epileptic seizures. The ESS predicts seizure based on patient's vital signs[4].

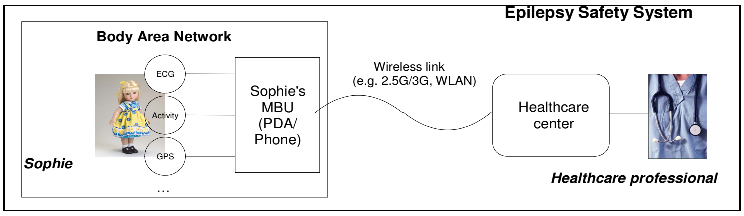

Figure 4. Epilepsy Safety System architecture

Therefore, Sophie wears a Body Area Network (BAN), responsible for vital signs' data collection and monitoring. The BAN consists of sensors (to measure ECG, activity), a GPS module (to determine her location) and a Mobile Base Unit (MBU). An internal ad-hoc communication network (e.g. Bluetooth) connects the sensors, GPS module and the MBU. The MBU is a gateway between the internal and external (2.5G/3G/WLAN) communication networks and can be used to (locally) process the vital signs data. The MBU can be implemented as a Personal Digital Assistant (PDA) or a mobile phone. Sophie's vital signs are in a reliable manner, real-time available to be viewed in the healthcare centre. This constitutes an *m-health tele-monitoring service*.
To make Sophie's data available to her healthcare professional, the MBU connects to one of the external communication networks (e.g., WLAN/2.5G/3G) as available at

[4] According to the clinical research, an epilepsy seizure can be predicted in 80% of the cases at best 30 seconds before it, based on ECG and an increasing heart rate of a patient [8].

her current location and time. To support Sophie's mobility, the MBU supports seamless handover between these networks.

Sophie's data have a maximum delay[5] defined by her doctor according to her current health state: 1 second for an emergency (i.e., seizure) case and 5 seconds otherwise. This requirement is defined such that the healthcare centre provides medical assistance to Sophie in time. Besides the delay, the doctor defined which basic data (e.g. alarm and location information in case of a seizure) must always be made available to him and which are redundant (e.g., ECG signal). The delay and basic/redundant data definitions constitute very important quality of service (QoS) requirements for the tele-monitoring service. The kind and volume of data send to the healthcare centre depends on the capabilities of (i.e. QoS offered by) networks available at Sophie's current location and time. The MBU always selects the external communication network depending on its offered QoS.

In case of a seizure, depending on its severity, the ESS alarms Sophie, her healthcare centre and eventually Sophie's mother (if strong seizure). These activities take in total few seconds and they constitute an *m-health alarm service*.

Sophie is biking to the library in her village. Although she feels good, the ESS warns her of a possible seizure and triggers an alarm at the healthcare centre. She stops biking and sits on a bench near-by. Before she can ask for help, a seizure starts.

The 3G network is available at Sophie's location, so all sensor data together with her location information are continuously sent to the healthcare centre. Based on the ECG signals her doctor sees, he decides to intervene. When the ambulance reaches Sophie, medical professionals provide her with medical assistance and take her to the hospital. Sophie's doctor continues monitoring her while she is being transported. During the ride, the ambulance moves out of the 3G network range and the MBU transparently connects to a 2.5G network. Once Sophie arrives at the hospital, the MBU connects to WLAN and her ECG signals are automatically displayed in the emergency room.

When the MBU switches between different communication networks, the ESS adapts the signals it sends to Sophie's doctor. As result, the doctor will not see Sophie's ECG signals when the ambulance moves out of the 3G network coverage and the 2.5G communication network is used.

5.3 Requirements for QoS- and context-aware mobile service infrastructure

To introduce QoS-awareness into a context-aware service infrastructure like AWARENESS, this infrastructure must meet the following requirements:

- QoS specification – the infrastructure must be able to handle user's end-to-end QoS specifications (e.g., data delay in our scenario).
- QoS mapping – the end-to-end QoS specifications must be mapped into the QoS requirements of the underlying communication network infrastructure.
- Assessment of the QoS-context information – the infrastructure must select the most suitable communication network based on the QoS offered by it and QoS actually required by a mobile user.
- QoS adaptation – when communication network handover occurs, the infrastructure must adapt service delivery to the QoS offered (section 5.5).

[5] time elapsed from the moment the data is gathered from Sophie's body to the moment it is displayed to her doctor in the healthcare centre

- Context adaptation – when context change occurs (e.g., an emergency), the QoS requirements may change. The infrastructure must incorporate context information when mapping QoS and adapting to QoS changes.
- QoS monitoring – the infrastructure must support real-time logging (i.e. measurements) of end-to-end QoS actually delivered to a mobile user.

5.4 Requirements for the QoS-context source

We indicate the QoS-context source as a source of the end-to-end QoS information aggregated over multiple mobile users. Following classification given in [10] we distinguish technology (i.e. network) oriented and user-oriented end-to-end QoS characteristics that this source aggregates. The network-oriented end-to-end QoS characteristic is its performance expressed in speed[6], accuracy[7] and dependability[8]. The user-oriented end-to-end QoS characteristics are a) perceived QoS, e.g., picture resolution, video rate/smoothness, audio quality, audio/video synchronization, b) service cost, e.g., per use or per unit cost and c) security level, e.g., authentication, confidentiality, integrity, non-repudiation. The requirements for a QoS-context source are following:

- speed – real-time calculations and response to a mobile user
- accuracy – degree of correctness with which QoS-context is provided by source
- dependability – source availability and reliability
- scalability – support for a number of mobile users
- QoS-context information management
  - QoS-context information aggregation, pre-processing and inference
  - QoS prediction e.g., along given user trajectory in particular timeframe

5.5 Incorporating context

To support our argument that the use of context in a mobile service infrastructure improves the delivered QoS, subsection 5.5.1 presents how application protocol adaptation can benefit from QoS-context information when transporting user data over the 3G wireless networks. Section 5.5.2 shows how location-specific QoS-context information can further improve QoS adaptation and delivered QoS.

5.5.1 Application protocol adaptation

The mobile service infrastructure acquires from the QoS-context source information about the QoS offered by communication network infrastructures (WLAN/2.5G/3G) available at the user's current location and time. Service delivery must be adapted to the currently offered QoS. To illustrate it we use our knowledge about QoS offered by 3G networks (details in [11, 12]). Figure 5 shows the delay and goodput[9] characteristics of a 3G network as a function of the application protocol's packet size sent in an uplink[10]. Larger packet size results in a higher goodput but higher delays.

In a context-unaware situation, the application protocol designer maps user's QoS requirements to some fixed application protocol packet size; for a low-delay, a small

---

[6] Time interval used to transport data from a source to a destination [2].
[7] The degree of correctness with which a service is performed [2].
[8] The degree of certainty with which the service can be used regardless of speed or accuracy [2].
[9] Goodput - a throughput of a communication network infrastructure observed at the application layer.
[10] from a mobile terminal to an application server

size could be chosen and for an efficient (and cost-effective) use of the 3G network a larger packet size could be chosen. If we incorporate context, this packet-size could be adapted at the service run-time according to the context-dependent user's QoS requirements.

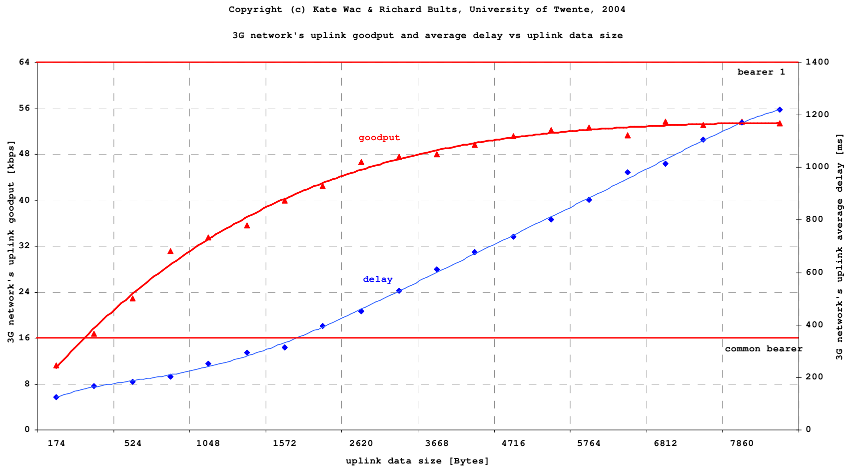

Figure 5. Performance characteristics of 3G communication network infrastructures

5.5.2 Location-and time-based QoS adaptation

Another way of utilizing the QoS-context information is conduction of a seamless handover to the communication network offering better QoS than the currently used network. As service user is mobile, underlying networks, and their offered QoS change along the user's trajectory. Offered QoS fluctuates over time because of changing number of mobile users at a given location and their changing demands. Ideally, QoS-context source provides the real-time QoS-context information for the service infrastructure and QoS predictions along the given trajectory and for particular timeframe. The infrastructure must (proactively) recognize a QoS-context change.

In a context-unaware situation, the system designer maps service user's QoS requirements into QoS offered by some default network. If we incorporate QoS-context information in the infrastructure, the network choice could be done at the service run-time according to the context-dependent user's QoS requirements.

6 Conclusions

In this paper, we discuss our ongoing research on QoS- and context-aware service infrastructure that supports the development of mobile applications in a heterogeneous network environment. We define a research problem and we indicate possible research trajectory towards a valid-able solution. Hence, our research trajectory aims to prove that the use of context information plays a significant role in required QoS specification and it improves the delivered QoS. We indicate the necessity of QoS-context source providing the service infrastructure with predictions on QoS offered by the networks at the service user's particular position and time. Due to the nature of the identified problem, our validation technique is system prototyping. We will validate if a mobile user experiences improvements in delivered QoS while using QoS- and context-aware service infrastructure, comparing to the QoS-unaware one, offering only the best-effort service.

We have already published one paper [13] disclosing our research topic and approving its relevance to the research community. We indicate the importance of our eventual findings and its applicability to improve QoS delivered by any mobile service. The target audience of our research is any mobile service provider, and as currently more and more applications and services go mobile, we indicate a high demand for our work. The beneficiaries of our work will be directly mobile users, experiencing their mobile services at the required quality. We identify the newness of our approach expanding beyond the standard QoS management framework comprising QoS contracts and QoS negotiation components. We propose user-driven approach, where user (and particularly a mobile application on user's behalf) will always have a choice amongst the underlying networks. This choice will be made with respect to user's QoS requirements.

Our research expands beyond the current telecom business model, where user is locked to one mobile operator. In our view, user needs to be able to make a decision, which network technology provided by which operator, is the most suitable to use. Moreover, by introducing the QoS-context source we further indicate user-empowerment; context information will be provided by users exclusively for users.

References

[1] D.Dey, *Providing Architectural Support for Context-Aware applications, PhD thesis, Georgia Institute of Technology*, USA, 2000

[2] ITU-T Recommendation I.350, *General aspects of Quality of Service and Network Performance in Digital Networks, including ISDNs*, March 1993

[3] Equanet project website, http://equanet.cs.utwente.nl

[4] CELLO project website, http://www.telecom.ntua.gr/cello

[5] J.Bolliger, T.Gross, *A framework based approach to the development of network aware applications*, IEEE Transactions on Software Engineering, vol. 24, no. 5, May 1998

[6] D.Chalmers, M.Sloman, *QoS and Context Awareness for Mobile Computing*, 1st Intl. Symposium on Handheld and Ubiquitous Computing (HUC'99), Germany, 1999

[7] J.Z.Sun, J.Sauvola, J.Riekki, *Application of Connectivity Information for Context Interpretation and Derivation*, 8th International Conference on Telecommunications (ConTEL 2005), Zagreb, Croatia

[8] M.Wegdam, *AWARENESS: A project on Context AWARE mobile NEtworks and ServiceS*, 14th Mobile & Wireless Communication Summit 2005, Germany

[9] I.Jacobson, G.Booch, J.Rumbaugh; *The Unified Software Development Process*; Addison-Wesley, 1999

[10] D.Chalmers, M.Sloman, *A Survey of Quality of Service in Mobile Computing Environments*, IEEE Communications Surveys, vol. 2, no. 2, 1999

[11] K.Wac, R.Bults et al., *Measurements-based performance evaluation of 3G wireless networks supporting m-health services*, Multimedia Computing and Networking 2005, Electronic Imaging Symposium, CA USA

[12] R.Bults, K.Wac et al. *Goodput analysis of 3G wireless networks supporting m-health services,* 8th International Conference on Telecommunications (ConTEL05), Croatia

[13] K.Wac, A.van Halteren, T.Broens, *Context-aware QoS provisioning in m-health service platform*, 11th Open European Summer School (EUNICE 2005), Colmenarejo, Spain.