

Dynamic Resources Allocation for Delivery of Personalized Services

Adam Grzech, Paweł Świątek, and Piotr Rygielski

Institute of Computer Science
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
{adam.grzech,pawel.swiatek,piotr.rygielski}@pwr.wroc.pl

Abstract. The aim of this paper is to introduce a problem of e-health services quality management. The process of delivering e-health services to users consists of two major tasks: service personalization and resources allocation. In this paper we introduce a use-cases of e-health system and distinguish services that can be offered. In order to satisfy user requirements one have to manage resources properly especially when communication conditions change (e.g. in an ambulance). An exemplary solution has been presented and conclusions for further work have been formulated.

Key words: e-Health, Service Oriented Architecture (SOA), service composition, quality of service (QoS), remote monitoring

1 Introduction

Advances in information and communication technologies allow health service providers to offer e-services in virtually every aspect of health care. The collection of e-health services facilitated by current and future ICT architectures ranges from remote health monitoring, through computer-aided diagnosis, video-consultation, health education, remote surgery and many others [9, 11].

Each of possible e-health services is a composition of atomic services provided by either ICT infrastructure or medical personnel. In some scenarios medical personnel can be both service provider and service consumer. Delivery of complex services as a composition of atomic services is the key feature of service oriented architecture (SOA) paradigm [10]. Application of SOA approach in e-health services delivery allows to personalize and flexibly adjust services to individual needs of service consumers.

It also allows to take into account changes in services environment; variable amount of resources implies necessity to reduce amount of services as well as quality of delivered services.

It is assumed that the set of available and required services is known. When the amount of available resources is sufficient all the ranked services may be delivered at assumed quality level. Decreasing amount of resources requires adaptation of services delivery scenarios. The gain of such an adaptation is to preserve

some services being important form the business process; list of preserved services depends on temporarily available resources. It is assumed that possible lists of services assure continuity of the business process.

According to the SOA paradigm each e-health service is composed of a set of atomic services providing required functionalities, which are delivered in certain predefined order. Each atomic service, which deliver certain functionality may be provided in different versions, which vary in non-functional characteristics such as: response time, security, availability, etc [10, 8]. Composition of complex services from different versions of atomic services allows to guarantee required quality of e-health services, which is very often critical in medical applications.

In this paper we present a general framework for QoS-aware SOA-based composition of e-health complex services in varying environment (e.g. varying communication channels capacities). Presented approach is explained with use of an illustrative example – remote monitoring of patients health parameters. The remote monitoring example includes: definition of the remote monitoring problem as a business process, identification of complex services playing major roles in the monitoring process, identification of the set of atomic services necessary to deliver required functionalities and composition of an exemplary complex service.

2 Remote monitoring of patients health

As an example of e-health business process consider the process of remote monitoring of patients health (see Fig. 1).

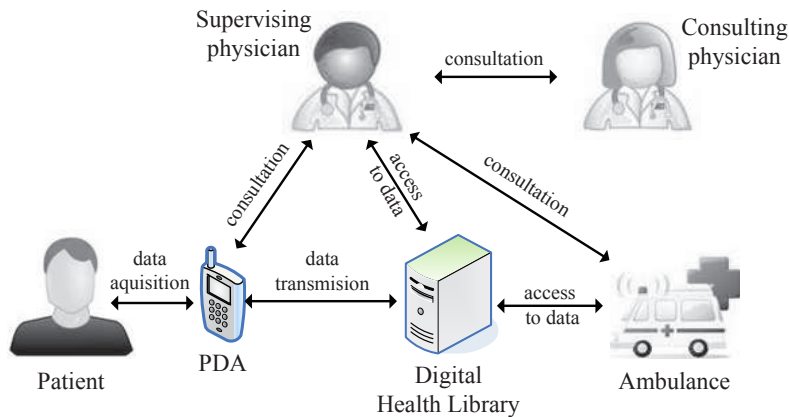


Fig. 1. Remote monitoring of patients health - an overview.

In this scenario it is assumed, that a patient is equipped with a mobile communication device (e.g. smartphone or PDA), which collects monitored data from

sensors placed on patients body. Collected data is preprocessed on patients mobile device and sent for further processing and storage to Digital Health Library. Further processing of collected data may involve among others: modelling and identification of physiological processes, updating of medical knowledge base, expansion of medical case study database and decision making for computer-assisted diagnosis. In the last case decision making algorithm may recognize any abnormal situations in patients health and notify medical personnel about such events.

An alarmed physician should have an access to patients health record, their current health state, medical knowledge bases and additional services such as consultation with patient or other physician. After gathering necessary information physician decides whether detected abnormal situation poses a threat to patients health and undertakes appropriate actions, e.g. sends an ambulance to the patient. In the life-threatening situation physician monitors patients status and instructs ambulance crew on how to proceed with a patient.

In the whole process of remote monitoring of patients health three complementary phases may be distinguished as presented on Fig. 2: basic monitoring, supervised monitoring and monitoring in the ambulance.

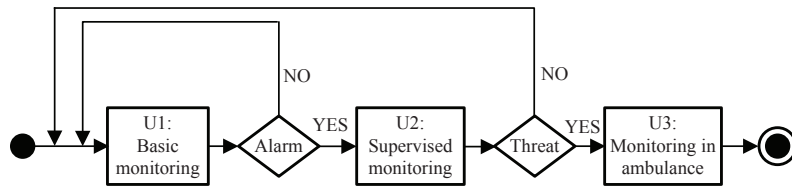


Fig. 2. The process of remote monitoring of patients health.

Each of presented monitoring phases is in fact a separate business process composed of one or more complex services. Basic monitoring covers routine day to day monitoring of patients vital parameters (e.g. ECG for post-cardiac patients). Supervised monitoring regards a case when abnormal behaviour of monitored parameters is detected. In this case an alarm is raised and other parameters (e.g.: body temperature, pulse and GPS coordinates) are monitored by physician providing him with information and services necessary to make decisions concerning further actions. Monitoring in ambulance applies to life-threatening situations in which an ambulance was dispatched to collect (possibly unconscious) patient. In this case both ambulance crew and physician should be provided with necessary information and communication services [9].

Processes of supervised monitoring and monitoring in ambulance can be decomposed into separate complex services. Such an exemplary decomposition of considered processes is presented on Fig. 3 a) and b) respectively.

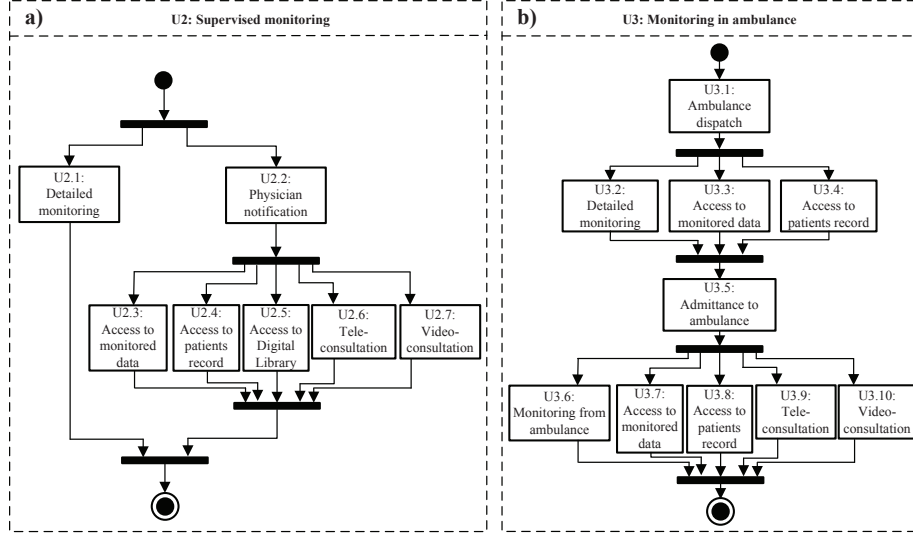


Fig. 3. The process of supervised monitoring (a) and the process of monitoring in an ambulance (b).

2.1 Supervised monitoring

The process of supervised monitoring starts at the moment of generation of an alarm concerning misbehaviour of patients health parameters. There are two immediate implications of such an alarm. The first one is the change in patients monitoring mode - more parameters are being recorded, possibly with denser time resolution. The second consequence is the notification of patients physician about possible health risks. From this moment the physician has access to personalised services allowing him to gather information about current state of the patient and make decision about further actions. These services may include (see Fig. 3 a)): on-line access to monitored health parameters (U2.3), access to patients health record (U2.4), access to Digital Health Library containing knowledge about similar cases (U2.5), and tele- and video-consultation with patient and/or medical experts (U2.6 and U2.7).

2.2 Monitoring in ambulance

The process of monitoring in an ambulance consists of two stages. The first one starts when supervising physician decides, that patients health is at risk and sends an ambulance to take the patient to the hospital (U3.1). From now on the crew of the ambulance has on-line access to monitored parameters of patients health (U3.3) as well as to the patients health record (U3.4).

After admittance to the ambulance (U3.5) patients monitoring mode changes again. All required patients health parameters are recorded now. Moreover,

audio-video monitoring of the patient in the ambulance is performed as well (U3.6). Besides previously accessible services (U3.7 and U3.8) ambulance crew can communicate with supervising physician through tele- and video-consultation services (U3.9 and U3.10).

The process of monitoring in an ambulance may be treated as personalized; it depends both on equipment available in the ambulance as well as on treatment procedures determined by ambulance crew.

3 Remote monitoring as business process

Business process is a series of interrelated activities or tasks that solve a particular problem or lead to achieve specific goal. In the SOA paradigm each of activities constituting in business process is represented as a complex service which delivers certain predefined functionality (see Fig. 4). Complex services, in turn, are composed of atomic services, which provide basic indivisible functionalities. The functionality of a complex service is an aggregation of functionalities of atomic services [5]. Similarly, the goal of a business process (its functionality) is an aggregation of functionalities of performed complex services.

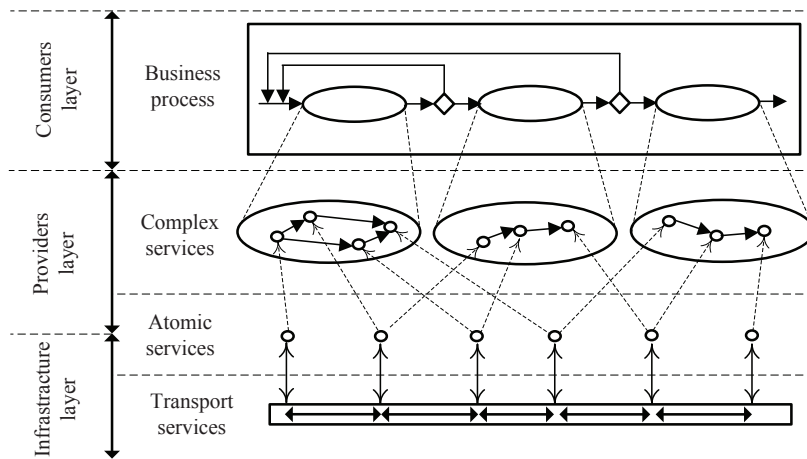


Fig. 4. Composition of business processes.

The difference between business process and complex service lies in that the former one is defined and composed by service consumer, while the latter one is delivered by service provider as a whole. Service consumer may influence the choice of particular complex services by specification of Service Level Agreement (SLA) containing functional and non-functional requirements. Service provider, on the other hand, composes complex services from available atomic services basing on requirements stated in the SLA [4].

In special cases the whole business process may be specified by single complex service. Such situations may occur for example when available processes are managed by single entity basing on certain regulations (e.g. medical processes in health care). In general, however, this approach is inefficient and inelastic, since it does not allow consumers to modify their requirements.

4 Services personalization and resources management

Depending on individual needs of service consumers and capabilities of execution environment each of available complex services can be delivered in many versions which differ in non-functional characteristics [2, 12]. As an example consider monitoring from ambulance service (U3.6) which allows to transmit monitored signals, voice and video from ambulance to supervising physician. Depending on individual case different health parameters are recorded and transmitted through the network (e.g.: ECG and pulse for cardiac patients or glucose for diabetics). Additionally voice and video transmission may be required by physician. Moreover individual requirements and amount of available system resources may influence the number and the quality of transmitted signals [6, 8].

Assume, that there are six signals possible to be measured and transmitted from ambulance to physician: video, voice, ECG, EMG, heart sound (HS) and heart rate (HR). Each signal can be measured and transmitted in two modes - high and low quality. High and low quality of health parameters can be reflected by higher and lower sampling rates. A transmission of high and low quality signals has different requirements for communication resources (see table 1).

Table 1. Throughput requirements for different signals and quality levels considered in the example.

Signal		Video	Voice	ECG	EMG	Heart sound	Heart rate
Required throughput [kbps]	High quality	5000	256	24	600	120	5
	Low quality	640	25	12	100	24	2

In general, preferences concerning the quality of requested complex service can be defined by penalty matrix $\mathbf{P}_l = [p_{ij}]$, where each element p_{ij} ($i = 1, \dots, I$; $j = 1, \dots, J$) represents a penalty for not delivering j -th signal in higher than i -th quality levels. The exemplary matrix for six signals considered above and three quality levels is defined by:

$$\mathbf{P}_l = \begin{bmatrix} p_{VID}^{HI} & p_{VOI}^{HI} & p_{ECG}^{HI} & p_{EMG}^{HI} & p_{HS}^{HI} & p_{HR}^{HI} \\ p_{VID}^{LO} & p_{VOI}^{LO} & p_{ECG}^{LO} & p_{EMG}^{LO} & p_{HS}^{LO} & p_{HR}^{LO} \\ p_{VID}^{NO} & p_{VOI}^{NO} & p_{ECG}^{NO} & p_{EMG}^{NO} & p_{HS}^{NO} & p_{HR}^{NO} \end{bmatrix}, \quad (1)$$

where for example $p_{21} = p_{VID}^{LO}$ is the penalty for not delivering high quality video signal and $p_{11} = p_{EMG}^{NO}$ is the penalty for not delivering EMG signal at all. If penalty $p_{ij} = 0$ then i -th quality level of j -th signal is not required.

Denote by \mathbf{D}_l a binary matrix of quality level delivery, where each element d_{ij} ($i = 1, \dots, I; j = 1, \dots, J$) is defined as follows:

$$d_{ij} = \begin{cases} 1 & i\text{-th quality level for } j\text{-th signal is delivered} \\ 0 & i\text{-th quality level for } j\text{-th signal is not delivered} \end{cases} \quad (2)$$

Exemplary matrix \mathbf{D}_l for a complex service in which video and voice signals are delivered at low quality, ECG signal is delivered at high quality, and remaining signals are not transmitted at all is presented below:

$$\mathbf{D}_l = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}. \quad (3)$$

Given penalty matrix \mathbf{P}_l and quality delivery matrix \mathbf{D}_l for certain complex service request req_l it is possible to calculate overall penalty p_l for not satisfying consumers preferences as follows:

$$p_l = \sum_{j=1}^J \mathbf{p}_{lj} \cdot \mathbf{d}_{lj}^T, \quad (4)$$

where \mathbf{p}_{lj} and \mathbf{d}_{lj} are j -th columns of matrices \mathbf{P}_l and \mathbf{D}_l respectively.

Let $\mathbf{R} = [r_{ij}]$ ($i = 1, \dots, I; j = 1, \dots, J$) denote the matrix of resources consumption, where each element r_{ij} represents the amount of resources required to deliver j -th signal at i -th quality level. In the example considered above resources requirements are stated in terms of required throughput (see table 1). Therefore the exemplary matrix \mathbf{R} is defined as follows:

$$\mathbf{R} = \begin{bmatrix} 5000 & 256 & 24 & 600 & 120 & 5 \\ 640 & 25 & 12 & 100 & 24 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (5)$$

Note, that elements of the last row of exemplary matrix \mathbf{R} are equal to zero since the lowest quality level represents situation in which signals are not delivered at all.

The amount of resources r_l necessary to deliver complex service at quality level represented by certain matrix \mathbf{D}_l can be calculated as follows:

$$r_l = \sum_{j=1}^J \mathbf{r}_j \cdot \mathbf{d}_{lj}^T, \quad (6)$$

where \mathbf{r}_j and \mathbf{d}_{lj} are j -th columns of matrices \mathbf{R} and \mathbf{D}_l respectively.

For the model presented above a number of resource management tasks can be formulated. The goal of each task may be different. For example one may want

to minimize the average or maximal penalty caused by violation of consumers individual preferences or to maximize consumers satisfaction for each incoming complex service request. Aforementioned tasks can be formulated as follows.

Task 1: Average penalty minimization

Given: set $L(t)$ of service requests currently being served, capacity C of the system and matrices of penalties \mathbf{P}_l and resources consumption \mathbf{R} .

Find: set of quality delivery matrices $\{\mathbf{D}_l^* : l \in L(t)\}$ such that average penalty caused by violation of consumers individual preferences is minimized:

$$\{\mathbf{D}_l^* : l \in L(t)\} = \arg \min_{\{\mathbf{D}_l : l \in L(t)\}} \sum_{l \in L(t)} \sum_{j=1}^J \mathbf{p}_{lj} \cdot \mathbf{d}_{lj}^T \quad (7)$$

with respect to system capacity constraints:

$$\sum_{l \in L(t)} \sum_{j=1}^J \mathbf{r}_j \cdot \mathbf{d}_{lj}^T \leq C. \quad (8)$$

Task 2: Maximal penalty minimization

This task is similar to Task 1 and can be derived by substitution of objective function in (7) by following formula:

$$\{\mathbf{D}_l^* : l \in L(t)\} = \arg \min_{\{\mathbf{D}_l : l \in L(t)\}} \max_{l \in L(t)} \sum_{j=1}^J \mathbf{p}_{lj} \cdot \mathbf{d}_{lj}^T \quad (9)$$

Task 3: Maximization of service consumer satisfaction

Given: the amount of resources $C(t)$ currently available in the system, matrix of preferences \mathbf{P}_l for incoming service request req_l and resources consumption matrix \mathbf{R} .

Find: composition of complex service defined by quality delivery matrix \mathbf{D}_l^* such that penalty caused by violation of consumers preferences is minimized:

$$\mathbf{D}_l^* = \arg \min_{\mathbf{D}_l} \sum_{j=1}^J \mathbf{p}_{lj} \cdot \mathbf{d}_{lj}^T \quad (10)$$

with respect to system capacity constraints:

$$\sum_{j=1}^J \mathbf{r}_j \cdot \mathbf{d}_{lj}^T \leq C(t). \quad (11)$$

The goal of Task 3 is to compose complex services according to consumers preferences. It can be performed each time when new complex service request arrives to the system. It can also be used for request admission control and Service Level Agreement renegotiation. Tasks 1 and 2, on the other hand, are performed in order to optimize utilization of systems resources and to guarantee average or minimal service consumer satisfaction.

4.1 Numerical example

Consider an example in which two monitoring service requests req_1 and req_2 arrive to the system. Requests req_1 and req_2 are characterized by following preferences matrices:

$$\mathbf{P}_1 = \begin{bmatrix} \times & 0 & \times & 0 & 0 & \times \\ \times & 100 & \times & \infty & 200 & \times \\ \times & \infty & \times & \infty & \infty & \times \end{bmatrix}, \mathbf{P}_2 = \begin{bmatrix} 0 & \times & \times & 0 & 0 & \times \\ 1000 & \times & \times & 100 & 10 & \times \\ \infty & \times & \times & \infty & 200 & \times \end{bmatrix} \quad (12)$$

which mean that request req_1 requires voice and heart sound signal to be delivered at least at low quality and EMG signal to be delivered at high quality. Similarly request req_2 requires video and EMG signal at least at low quality and additional heart sound signal would improve consumers satisfaction.

Assume, that systems capacity is equal to $C = 2510kbps$ and that capacity requirements of available signals are given by matrix \mathbf{R} defined in (5). As a result of minimizing average penalty for not satisfying consumers preferences (Task 1 defined by (7) and (8)) following quality delivery matrices are calculated:

$$\mathbf{D}_1 = \begin{bmatrix} \times & 1 & \times & 1 & 1 & \times \\ \times & 0 & \times & 0 & 0 & \times \\ \times & 0 & \times & 0 & 0 & \times \end{bmatrix}, \mathbf{D}_2 = \begin{bmatrix} 0 & \times & \times & 1 & 1 & \times \\ 1 & \times & \times & 0 & 0 & \times \\ 0 & \times & \times & 0 & 0 & \times \end{bmatrix} \quad (13)$$

resulting in overall penalty $p = 1000$ for not delivering high quality video signal for second service request req_2 . Delivery of high quality video signal in this example is impossible because HQ video capacity requirements are higher than overall capacity C of the system. Services composition and resources allocation represented by matrices \mathbf{D}_1 and \mathbf{D}_2 (13) are illustrated on Fig. 5 (state before moment t_1).

Assume, that at certain moment t_1 a third service request req_3 characterized by matrix \mathbf{P}_3 arrives to the system. In order minimize average penalty for not satisfying consumers preferences systems resources have to be reallocated. Final service composition and resources allocation depends on penalty matrices \mathbf{P}_1 , \mathbf{P}_2 and \mathbf{P}_3 .

In order to show the difference between final resources allocation assume two alternate matrices \mathbf{P}_3 :

$$\mathbf{P}_3^a = \begin{bmatrix} 0 & 0 & \times & 0 & 0 & \times \\ 1000 & \mathbf{250} & \times & \infty & \infty & \times \\ \infty & \infty & \times & \infty & \infty & \times \end{bmatrix}, \mathbf{P}_3^b = \begin{bmatrix} 0 & 0 & \times & 0 & 0 & \times \\ 1000 & \mathbf{150} & \times & \infty & \infty & \times \\ \infty & \infty & \times & \infty & \infty & \times \end{bmatrix} \quad (14)$$

which differ in the penalty for not delivering high quality voice signal. Two different solutions of resources allocation task are represented by quality level matrices $\mathbf{D}_1^a, \mathbf{D}_2^a, \mathbf{D}_3^a$ and $\mathbf{D}_1^b, \mathbf{D}_2^b, \mathbf{D}_3^b$:

$$\mathbf{D}_1^a = \begin{bmatrix} \times & 0 & \times & 1 & 0 & \times \\ \times & 1 & \times & 0 & 1 & \times \\ \times & 0 & \times & 0 & 0 & \times \end{bmatrix}, \mathbf{D}_2^a = \begin{bmatrix} 0 & \times & \times & 0 & 0 & \times \\ 1 & \times & \times & 1 & 1 & \times \\ 0 & \times & \times & 0 & 0 & \times \end{bmatrix}, \mathbf{D}_3^a = \begin{bmatrix} 0 & 1 & \times & 1 & 1 & \times \\ 1 & 0 & \times & 0 & 0 & \times \\ 0 & 0 & \times & 0 & 0 & \times \end{bmatrix} \quad (15)$$

and

$$\mathbf{D}_1^b = \begin{bmatrix} \times & 0 & \times & 1 & 1 & \times \\ \times & 1 & \times & 0 & 0 & \times \\ \times & 0 & \times & 0 & 0 & \times \end{bmatrix}, \mathbf{D}_2^b = \begin{bmatrix} 0 & \times & \times & 0 & 1 & \times \\ 1 & \times & \times & 1 & 0 & \times \\ 0 & \times & \times & 0 & 0 & \times \end{bmatrix}, \mathbf{D}_3^b = \begin{bmatrix} 0 & 0 & \times & 1 & 1 & \times \\ 1 & 1 & \times & 0 & 0 & \times \\ 0 & 0 & \times & 0 & 0 & \times \end{bmatrix}. \quad (16)$$

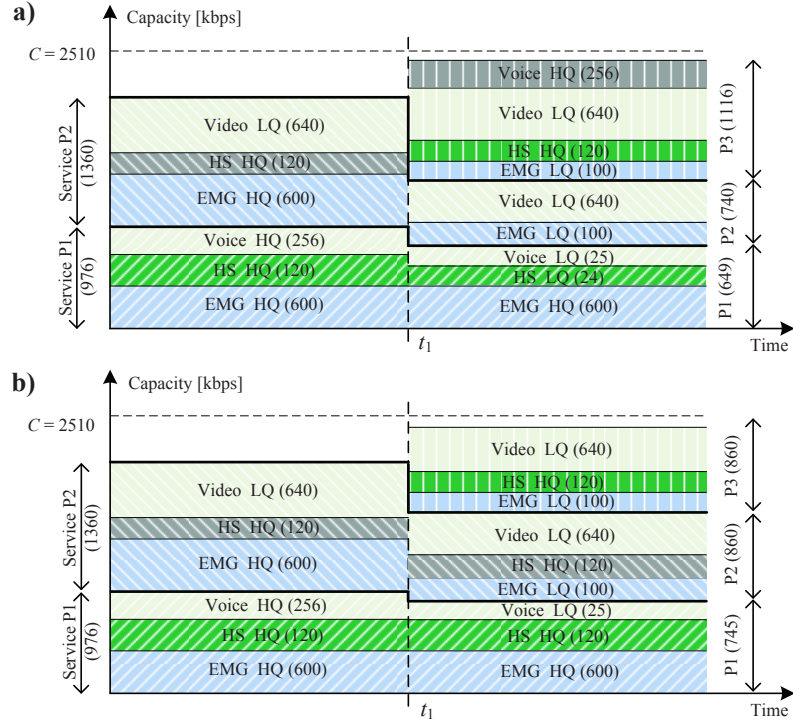


Fig. 5. Services composition and resources reallocation for two versions of third request preferences differing in the value of penalty p_{322} for not delivering HQ voice for request req_3 : (a) $p_{322} = 250$, (b) $p_{322} = 150$.

Final services composition and resources allocation represented by above matrices are illustrated on figures Fig. 5a) and Fig. 5b) respectively. Note that in

the presented example, depending on the penalty for not delivering HQ voice for request req_3 , HQ voice in req_3 is traded for HQ heart sound in req_1 and req_2 . The threshold value of this penalty is $p_{322} = p_{125} + p_{225} = 210$.

5 Conclusions

In this paper a general problem of e-health services management was introduced. It consists of two major tasks: service personalization and resources allocation. Service personalization allows to flexibly adjust delivered services based on individual needs of service consumers. Resources management allows to reserve and allocate resources necessary to deliver requested services and satisfy consumers preferences. In the presented approach both tasks of personalization and resource allocation are solved simultaneously as single optimization problem in which certain parameters concern the personalization task (penalty matrices), while other parameters regard allocation tasks (resources consumption matrix). Unfortunately formulated tasks are in general NP-hard [3], therefore heuristic algorithms should be applied to effectively control processes of service composition, personalization and resources management [1, 7].

Presented approach allows to flexibly compose complex remote monitoring services from a given set of atomic communication services. Service personalization means that complex service components (atomic services) are given certain priorities based on which conflicts are resolved in the case of insufficient resources. Such personalization allows to deliver services with lower quality which still meet consumers requirements.

Reduction of the quality of complex service means deterioration of the quality of atomic service components and/or decrease in the number of utilized service components. Reduction of the quality or the number of service components is performed based on given or acquired knowledge about medical procedures. Acquired knowledge may also address personal experiences, expertise and preferences of systems consumers.

Acknowledgements

The research presented in this paper has been partially supported by the European Union within the European Regional Development Fund program no. POIG.01.03.01-00-008/08.

References

1. Albrecht Jeannie, Oppenheimer David, Vahdat Amin and Patterson David A., "Design and Implementation Trade-offs for Wide-area Resource Discovery", ACM Transactions on Internet Technology, Vol. 8 , Issue 4, Article 18, September 2008.

2. M. Alrifai and T. Risse, "Combining global optimization with local selection for efficient QoS-aware service composition", WWW '09: Proceedings of the 18th international conference on World wide web, pp.881–890, ACM 2009.
3. Garey Michael R. and Johnson David S., *Computers and Intractability: A Guide to the Theory of NP-completeness*, Publ. W. H. Freeman, 1979.
4. Grzech A., Rygielski P. and Świątek, P.: QoS-aware infrastructure resources allocation in systems based on service-oriented architecture paradigm. HET-NETs 2010, pp. 35-48, Zakopane, Poland, 2010.
5. Grzech A. and Świątek P., "Modeling and optimization of complex services in service-based systems" *Cybernetics and Systems*. 2009, vol. 40, pp. 706-723
6. Grzech A. and Świątek P., "Parallel processing of connection streams in nodes of packet-switched computer communication systems". *Cybernetics and Systems*. 2008, vol. 39, nr 2, pp. 155-170
7. Karlsson Magnus, "Design Rules for Producing Controllable Computer Services", Proc. of 10th IEEE/IFIP Network Operations and Management Symposium, pp. 1-14, 2006.
8. Khan Shahadat, Li Kin F. and Manning Eric G., "The Utility Model For Adaptive Multimedia Systems" In *International Conference on Multimedia Modeling*, 1997 pp.111-126.
9. Niyato D., Hossain E., and J. Diamond, "IEEE 802.16/WiMax-based broadband wireless access and its application for telemedicine/e-health services", *IEEE Wireless Communications*, vol. 14, no. 1, 2007, pp. 72-83.
10. Rygielski P., Świątek P., "QoS-aware Complex Service Composition in SOA-based Systems", in "SOA Infrastructure Tools: Concepts and Methods", Springer-Verlag, Berlin 2010.
11. Vergados D.J., Vergados, D.D. and Maglogiannis, I. , "NGL03-6: Applying Wireless DiffServ for QoS Provisioning in Mobile Emergency Telemedicine," *Global Telecommunications Conference*, 2006. GLOBECOM '06. IEEE, pp.1-5, Nov. 27 2006-Dec. 1 2006.
12. Yu Tao, Zhang Yue and Kwei-Jay Lin, "Efficient algorithms for Web services selection with end-to-end QoS constraints", *ACM Trans. Web* 2007, Vol. 1, No. 1.