

The Influence of SIP Call Control Signalling on VoIP Quality of Experience

Jasmina BARAKOVIĆ HUSIĆ, Sabina BARAKOVIĆ, Seudin KASUMOVIĆ

Abstract: The rapid growth in subscribers and usage of multimedia services enlarges the volume of Session Initiation Protocol (SIP) call control signalling creating a need to understand Quality of Experience (QoE) in this case and improve it. This paper provides an analysis of influence of SIP call control signalling on QoE for Voice over Internet Protocol (VoIP) service. The aim was to investigate whether SIP call control signalling load has the influence on the human perception of SIP signalling performances and QoE, and to identify the importance of distinct SIP-based signalling performance metrics. Moreover, the intention was to determine whether SIP call control signalling load changes its impact if previously proposed algorithm for differentiated treatment of SIP messages is activated, and quantify mutual relationships of considered user perceptions and QoE. The findings show that SIP call control signalling load has a strong and negative impact on dependent variables and that the proposed algorithm improves QoE and human perception of SIP signalling performances.

Keywords: QoE; SIP call control signalling; SIP performance; VoIP

1 INTRODUCTION

Network transformation towards the fifth generation (5G) and modern devices production have allowed the use of a large number of attractive services. They are offered by service providers in order to balance falling revenues from standard voice and high-speed, flat-rate services. Those are multimedia services, which usually include voice, video, instant messaging and presence. However, nowadays humans are quality meters, and their requirements, perceptions and experiences considering a particular service carry higher value [1]. In extremely competitive conditions, users have the right to choose from plenty of service providers. Thus, it is not enough anymore to simply make the service available to users. Those services have to be delivered in such a way that users experience high quality at acceptable price.

A lot of services that are running as data traffic need a high service quality in order to satisfy user's requirements and expectations. An integral part of these services is the Session Initiation Protocol (SIP) signalling. In this regard, Voice over Internet Protocol (VoIP) session can be decomposed into the following components: registration, call setup, voice quality, and call teardown [2]. Traditional service quality considerations refer to media service components. However, human perception of service quality depends on both media and signalling service components, since long call setup delay or low voice quality results in bad user experience. Furthermore, SIP signalling is in charge of service quality negotiation [3]. Consequently, signalling service components should be given greater importance since they can be used for prevention of the service quality degradation.

Service providers are faced with the issue of fast growing volume of SIP call control signalling. Therefore, they have to set up mechanisms for controlling increasing amount of this type of signalling produced by various communication services for enhancing Quality of Experience (QoE) [4]. In order to improve service and experience quality, we proposed the algorithm for differentiated treatment of SIP messages. General idea behind this algorithm is proposed in our previous work [5], whereas its implementation in Network Simulator version 2 (ns-2) is described in [6]. Encouraged by the results obtained in simulation environment, we have developed and deployed the algorithm for experimental usage in form

of SIPPRIO package [7]. The algorithm is verified through research study of SIP call control signalling procedures under different conditions. Summarizing our research findings, the open issue is identified and covered in this paper. It is focused on the investigation of relationship between SIP call control signalling and QoE, and whether the algorithm for differentiated treatment of SIP messages influences that relationship.

The paper has the following structure: after the introduction, Section 2 gives review of related work in the area of SIP signalling influence on service quality and QoE. Section 3 provides the objectives and hypotheses and explains the research methodology. Furthermore, this section describes the research steps performed for modelling and managing the relationship between SIP call control signalling and QoE. The first step includes designing, setting up and conducting experiments needed to obtain data for describing the interaction between SIP call control signalling and QoE, and the influence of the proposed algorithm for differentiated treatment of SIP messages. Therefore, a relevant number of participants tested and subjectively evaluated the VoIP service by using developed evaluation surveys in order to derive statistically and practically significant results. In Section 4, the obtained results are discussed and statistically processed using descriptive statistics, Analysis of Variance (ANOVA), and correlation and regression analysis. The analysis results are used to draw conclusions and define activities for future work that are given in Section 5.

2 RELATED WORK

This section gives a brief overview of related work considering signalling performance for VoIP service. According to literature survey summarized in [8], it can be noticed that signalling performance has not been investigated as much as media performance. Reviewed research studies consider the influence of call control signalling on the service quality including measurement and evaluation of SIP performance metrics. The growing use of SIP signalling in these services has led to creation of measurement methodology for SIP server performances [9], [10]. Many research activities have been directed to evaluation of various SIP performance metrics in different environments, e.g., Mobile Ad Hoc Network (MANET) [11], Long Term Evolution (LTE) [12], [13], IP

Multimedia Subsystem (IMS) [14], Software Defined Network (SDN) [15]/ Network Function Virtualization (NFV) environment [16], [17].

Various signalling performance metrics were defined by two standardization organizations: International Telecommunication Union (ITU) and Internet Engineering Task Force (IETF). IETF introduced a set of SIP signalling metrics in Request for Comments (RFC) 6076 [18]. In this paper, we discuss the Registration Request Delay (RRD), Session Request Delay (SRD), and Session Disconnect Delay (SDD). Almost identical measures were determined by ITU. Recommendation E.721 defines a metric called Post Selection Delay (PSD), which is similar to SRD, while Recommendation E.411 defines Network Effectiveness Ratio (NER) and Answer Seizure Ratio (ASR), which are similar to Session Establishment Ratio (SER) and Session Establishment Effectiveness Ratio (SEER), as defined in RFC 6076. These standardization activities led to the research that resulted in the proposal of new signalling performance metrics.

Based on the literature review in [19], one may note that Session Setup Delay (SSD) was usually analysed as SIP performance metric. This metric is comparable to Call Setup Delay (CSD) that has been described in [20]. Since these signalling performances were not satisfactory to determine the signals exchanged through the whole session, a novel concept called Quality of Signalling (QoS_g) was described in [21]. It introduced the following metrics: Response Delay (RpD), Processing Delay (PD), and User to User Delay (UUD). In order to precisely evaluate the process of session negotiation, Session Renegotiation Delay (SRNT) and Session Negotiation Time (SNT) were presented in [22]. In addition, the four metrics that quantify the overhead associated with IMS were defined in [23] as follows: Registration Time (RT), Initial Response Time (IST), Initial Ringing Time (IRT), and Disconnect Request Time (DRT). Signalling performance of next generation network system (i.e., EuQoS system) was studied in [24].

Considering this brief review of related literature and studies, one can note that they are directed to analysing the influence of call control signalling on Quality of Service (QoS). Nevertheless, there is no guarantee that acceptable QoS means acceptable QoE for end user. Accordingly, there is a need to investigate the interaction between call control signalling and QoE and human perception of signalling performance. Therefore, the research study described in Section 3 is performed analysing aforementioned interaction in two scenarios: one which identifies whether the influence of SIP call control signalling load on QoE exists, and other dealing with the benefits that the activation of the algorithm for differentiated treatment of SIP messages, described in detail in [7], could have on user's QoE.

3 RESEARCH METHODOLOGY

3.1 Objectives and Hypotheses

Based on the related work and reviewed literature, we drew ten hypotheses organized in four groups: (i) influence of call control signalling load on human perception of SIP signalling performances and QoE (hypotheses H1.1-H1.4), (ii) finding the relationship between human perception of SIP signalling performances and QoE (hypothesis H2), (iii) influence of call control signalling load on human

perception of SIP signalling performances and QoE when the proposed algorithm for differentiated treatment of SIP messages is activated (hypotheses H3.1-H3.4), and (iv) finding the relationship between human perception of signalling performances and QoE when the proposed algorithm for differentiated treatment of SIP messages is activated (hypothesis H4).

The first group of hypotheses attempts to highlight the existence of influence of call control signalling load on human perception of SIP signalling performances and QoE. Our interest is to define the extent to which the analysed influence factor affects the considered dependent variables. The expectation is to discover a strong and negative influence of call control signalling load, meaning that if the load is higher, then human perception of considered SIP signalling performances and QoE will be lower.

H1.1: The increase of SIP call control signalling load has strong and negative influence on the QoE.

H1.2: The increase of SIP call control signalling load has strong and negative influence on the human perception of RRD.

H1.3: The increase of SIP call control signalling load has strong and negative influence on the human perception of SRD.

H1.4: The increase of SIP call control signalling load has strong and negative influence on the human perception of SDD.

Although state-of-the-art literature review addresses the relationships of QoE in terms of VoIP with many other relevant influence factors, as stated in the related work section, there is no prior work considering VoIP QoE in terms of perception of SIP signalling performances. Therefore, in second hypothesis H2 we expect to obtain model that will quantify mutual relationships of analysed human perceptions and QoE and allow us to identify the importance of distinct call control signalling performances.

H2: The QoE for VoIP is affected by the human perceptions of SIP signalling performances following the sequence from most to least important: human perception of SRD, human perception of SDD, and human perception of RRD.

Since the rapidly increasing amount of call control signalling generated by different types of communication services affects the service quality and potentially the user QoE with VoIP, we proposed the algorithm for differentiated treatment of SIP messages. This algorithm is developed for a SIP call control server and differs from conventional First In First Out (FIFO) scheduling by classifying packets containing SIP message into three priority queues. The general idea of this algorithm is to assign higher priority values to packets containing SIP message that terminate the call control session. This may serve for reducing the SIP server usage, and by that means making better the QoS and QoE. On the contrary, giving lower priority to packets containing SIP message that establish call control session may refuse the session, which can enhance the QoS and QoE.

Therefore, the third group of hypotheses attempts to highlight if the algorithm for differentiated treatment of SIP messages influences human perception of SIP signalling performances and QoE for VoIP service. Our interest in defining the extent to which this algorithm

influences considered dependent variables. It is expected to discover a less strong but still negative influence of SIP signalling load, meaning that if the load is higher, then human perception of SIP signalling performances and QoE will be lower, but less.

H3.1: The proposed algorithm for differentiated treatment of SIP messages decreases strong effect of signalling load increase on overall user satisfaction.

H3.2: The proposed signalling algorithm decreases strong effect of signalling load increase on user perception of *RRD*.

H3.3: The proposed signalling algorithm decreases strong effect of signalling load increase on user perception of *SRD*.

H3.4: The proposed signalling algorithm decreases strong effect of signalling load increase on user perception of *SDD*.

Furthermore, in the fourth hypothesis H4 we look for obtaining a model that will quantify mutual relationships of analysed human perceptions and *QoE* when the proposed algorithm is activated and enables the identification of importance of different SIP signalling performances.

H4: The overall user *QoE* with VoIP when the proposed algorithm is used is affected by the human perceptions of SIP signalling performances following the sequence from most to least important: human perception of *SRD*, human perception of *SDD*, and human perception of *RRD*.

3.2 Experimental Environment and Organization

For the sake of obtaining data describing the relationship between call control signalling and user satisfaction, an experimental environment is established as follows. All software tools are installed on Toshiba Satellite S875 laptop, which has the following specifications: Intel Core i7-3630QM processor (6M Cache, up to 3,4 GHz), 8 GB RAM (1600 MHz), 1 TB hard disk drive, AMD Radeon 7670M graphics card. This laptop runs Windows 8.1 operating system. Oracle VM Virtual Box is used for creation of eight virtual machines, running two different Linux releases, i.e. 12.04 LTS and 13.04. Three virtual machines are used to run SIPp, a traffic generator for SIP, and Low Orbit Ion Cannon (LOIC), an application for network stress testing. Other four virtual machines serve as a basis to run Jitsi, a freely available multi platform voice, video conferencing and messaging application. This application is used to test user satisfaction with register, call setup and teardown procedures. Remaining virtual machine serves as a basis to place the Kamailio SIP server, which is integrated with the SIPPRIO package developed for the purpose of SIP message differentiation.

Several experiments are designed, setup, and conducted to investigate the influence of SIP call control signalling load and messages differentiation on the *QoE* and human perception of signalling performances. Each experiment is organized as a combination of two scenarios. These scenarios differ from each other in whether or not the algorithm for differentiated treatment of SIP messages is activated. Scenario 1 means that the algorithm is inactive, whereas Scenario 2 indicates that the algorithm is made

active via user-space application by setting queue weights as follows: Q1 weight: 4, Q2 weight: 3, Q3 weight: 2, and Q4 weight: 1 [7].

The SIP performance metrics are measured in both scenarios during VoIP signalling procedures (registration, call setup and teardown procedure). Therefore, the focus is on *RRD*, *SRD*, and *SDD*, which are introduced in RFC 6076. These SIP signalling performances are collected in nine distinct load conditions.

Various types and amount of SIP messages burden the Kamailio SIP server. The load structure is taken from the British Telecom model of signalling traffic [25]. According to this model, the number of major traffic flow instances per second in busy hour has been calculated for 2 million subscribers. These traffic flows were chosen since they were considered likely to provide the bulk of the traffic. The number of traffic flows instances per second is used to determine the number of associated SIP messages. The number of corresponding SIP messages is multiplied by the typical SIP message size taken at a lightly loaded Kamailio SIP server interface. These multiplications determine the signalling load size and structure that forms a basis for nine different measurement points. Signalling load increases from zero value in first measurement point to maximum value that equals 5,86 MB in ninth measurement point.

30 participants of different age and gender groups have been involved in the realization of experiments. The study sample includes 14 females (47%) and 16 males (53%), belonging to one of five age groups. The mean age of all participants is 28,63. 4 participants (13%) of age less than 15 years old; 6 participants (20%) age 15-18 years old; 5 participants (17%) age 19-23 years old; 9 participants (30%) age 24-41 years old; 6 participants (20%) age more than 41 years old. All participants are given a task of registering, establishing, and terminating the one VoIP call in both scenarios, resulting in 72 tasks per participant, and 2160 tasks in the experimental study. After executing each scenario, they are requested to evaluate their overall satisfaction when VoIP is used and rate the signalling service components (e.g., Evaluate the *RRD/SRD/SDD* while using VoIP).

The subjective assessment of test VoIP calls is conducted using the evaluation survey. Acquisition of data is performed electronically and through evaluation form which has two parts. The first part of the survey consists of the seven queries covering participant's private information and prior experience with using the VoIP. Collected data reveal that 30% of participants use the VoIP service on a daily basis, 30% every two-three days, while 10% and 3% access the VoIP service once per week and month, respectively. 27% of participants never use the VoIP service.

The second part covers participant's ratings in relation to offered statements. Four statements refer to overall user satisfaction and satisfaction with register, setup and teardown procedures while using VoIP service. The first section of the second part of the survey contains statement related to overall user satisfaction with VoIP service. This statement is Mean Opinion Score (MOS) specified in ITU-T Recommendation P.800.1 that is most commonly used in *QoE* studies [1]. The second section of the second part of the survey refers to evaluation of user satisfaction with

register, setup and teardown procedures while using VoIP service. Three statements are offered and evaluated on a 5-point Likert scale, where 1 means "strongly disagree", 2 "disagree", 3 "neutral", 4 "agree", and 5 "strongly agree".

The experiment procedure includes the three following steps [1]: introduction and explanation of the tasks to be done by the participants, participant training, and iterative testing and evaluating of experimental VoIP calls. The first step lasts 8 minutes and introduces the research topic and assessment ratings. In addition, prior VoIP call testing, participants are instructed what actions to perform. The explanation is followed with a training session that is necessary to provide accurate task performance and practice the use of scale to rate the quality and satisfaction. The duration of the second step is 5 minutes. The final step includes testing and evaluating, and lasts approximately 1,5 hours for each participant. The participants (i.e., family members, colleagues and friends) were instructed to be intuitive rather than to think about how they feel. Experiments were carried out at authors' homes (in Sarajevo, Tuzla, and Zenica) with previously mentioned facilities [1].

4 RESEARCH RESULTS AND DISCUSSION

To test the first group of hypotheses (H1.1-H1.4) stating that distinctions in the levels of signalling load affect human perception of SIP signalling performances and *QoE* when using VoIP, we ran one-way ANOVA whose results are presented in Tab. 1. Independent variable (IV) in this statistical test is signalling load, while dependent variables (DVs) are human perceptions of SIP signalling performances and overall *QoE*. Based on given numbers, one may conclude that there exists strong ($p < 0,001$) statistically and practically (η^2) significant influence of signalling load on human perception of SIP signalling performances, and overall *QoE*. The results show that the impact is negative as well, given the fact that the higher the signalling load is, the lower user rates for DVs are (means). This implies that hypotheses H1.1-H1.4 are supported.

Further on, in order to additionally explain the relation between considered DVs and IV, and reveal if the variation of human perceptions of *RRD*, *SRD*, and *SDD*, and overall *QoE* is explained by the variation of call control signalling load, a regression analysis was performed. The resulting four models (for each of DVs), which are provided in Tab. 1, show cubic and quadratic (human perception of *SRD*) nature. The percent of variability of *QoE* can be explained by knowing call control signalling load by approximately 90,6%, whereas the coefficient of variation (*CV*) is about 16,8%. Further on, the percent of variation of human perception of *RRD*, *SRD* and *SDD* can be accounted for by the knowledge of signalling load by approximately 84%, 83,3% and 86,7%, while the calculated *CV* is approximately 16,2%, 26,6% and 18,9%, respectively. These results indicate that the models given in Tab. 1 represent quantification of mutual relationships of signalling load and human perceptions of SIP signalling performances and *QoE* for VoIP. The models characterized by such a high R^2 factor are outstanding, since it is not easy to find model that can describe human behaviour with such high explaining ability. Additionally, obtained *CV* for each

of the cases shows that conducted analysis may be used for the predicting intent. The second aim of this paper was to quantify the relationships between human perception of SIP signalling performances and *QoE*. In other words, in order to test hypothesis H2 and to investigate how each of these human perceptions is connected to *QoE* and how it makes contributions to it, multiple linear regression analysis has been conducted with three predictors being the considered human perceptions. The resulting linear model is given by Eq. (1). The percent of variability of *QoE* can be explained by knowing human perception of *SRD* and *SDD* by approximately 85,1%, whereas the obtained coefficient of variation is about 21,2%.

$$QoE = 0,234 + 0,503 \cdot UPSRD + 0,421 \cdot UPSDD. \quad (1)$$

According to aforementioned results, it can be concluded that the proposed model represents quantification of mutual relationships of human perceptions of SIP signalling performances and *QoE* for VoIP, since, as indicated earlier, models characterized by high R^2 factor are outstanding. In accordance with this model, we determined the importance of different dimensions in terms of analysed human perceptions and *QoE*. Namely, human perception of *SRD* and *SDD* varies to extent which they affect the *QoE* (going from most to least important), while human perception of *RRD* has no influence of the *QoE*. This implies that assumed order of importance in hypothesis H2 is not supported. This result may be explained by the fact that user is mainly focused on *SRD* and *SDD*, while *RRD* is not perceived as a service component. With this in mind, service providers may use the obtained linear regression model for proposing and developing new commercial service models. Namely, it may be used to improve the charging user experience, since most of service providers charge the session setup. Since overall *QoE* with VoIP and *SRD* are strongly negatively correlated, the service providers may use this regression model to decide depending on context whether or not the call setup should be charged. Moreover, it may be used to predict the acceptable values of *SDD* that do not degrade the overall *QoE*, because lower *SDD* may be used to enhance the charging user experience.

Additional aim of the paper was to validate the algorithm for differentiated treatment of SIP messages in *QoE* context and test the third group of hypotheses. For those purposes, we have conducted one- and two-way ANOVA analysis, whose outcomes are given in Tab. 2 and Tab. 3, and graphically represented in Fig. 1. Based on given numbers, one may conclude that there exists strong ($p < 0,001$) statistically and practically (η^2) significant influence of signalling load on human perception of SIP signalling performances, and *QoE* when the algorithm is activated. This was expected as well as the negative impact, given the fact that the higher the signalling load is, the lower user rates for DVs are (means). However, opposite to the means per each signalling load obtained for scenario where the relations were examined without the implementation of the algorithm, the values in this case are quite higher indicating that the algorithm decreases the effect of call control signalling load on considered variables. This is confirmed with the results of two-way ANOVA analysis, which indicate that there exists

statistically ($p < 0,001$) and practically (η^2) significant interaction between the effects of the algorithm usage and signalling load level on human perceptions of SIP signalling performances and overall *QoE*. Thereby, one may conclude that hypotheses H3.1-H3.4 are supported.

Further on, a regression analysis was performed with the purpose of additional explanation of the relationship

between analysed *DVs* and *IV* when the algorithm is used and reveal how the variability of human perceptions of *RRD*, *SRD*, and *SDD*, and overall user *QoE* is explained by the variability of signalling load. The resulting four models (for each of *DVs*) are provided in Tab. 2 and show cubic (user perception of *SRD*) and quadratic nature.

Table 1 Summary of results for influences of signalling load on SIP call control signalling performances and *QoE*

Signalling load (MB)	Overall user <i>QoE</i> (<i>M</i> (<i>SD</i>))	Human perception of <i>RRD</i> (<i>M</i> (<i>SD</i>))	Human perception of <i>SRD</i> (<i>M</i> (<i>SD</i>))	Human perception of <i>SDD</i> (<i>M</i> (<i>SD</i>))
0,00	5,00 (0)	5,00 (0)	5,00 (0)	5,00 (0)
0,73	5,00 (0)	5,00 (0)	5,00 (0)	5,00 (0)
1,47	4,90 (0,305)	4,93 (0,254)	4,80 (0,407)	5,00 (0)
2,20	4,23 (0,679)	4,47 (0,571)	3,47 (0,860)	5,00 (0)
2,93	2,83 (0,913)	3,93 (0,785)	1,77 (0,935)	4,47 (0,571)
3,66	2,53 (0,730)	3,73 (0,583)	1,67 (0,884)	4,30 (0,596)
4,40	1,33 (0,479)	2,70 (0,750)	1,00 (0)	1,00 (0)
5,13	1,03 (0,183)	1,50 (0,731)	1,00 (0)	1,00 (0)
5,86	1,03 (0,183)	1,37 (0,615)	1,00 (0)	1,00 (0)
<i>F</i> (8, 270)	362,274	200,688	1451,614	344,721
<i>p</i>	<0,001	<0,001	<0,001	<0,001
η^2	0,917	0,860	0,978	0,914
Regression model	$4,955 + 0,703 \times SL^3 - 0,660 \times SL^2 + 0,073 \times SL$	$4,951 + 0,352 \times SL^3 - 0,286 \times SL^2 + 0,020 \times SL$	$5,650 + 0,082 \times SL^2 - 1,333 \times SL$	$4,685 + 1,437 \times SL^3 - 0,743 \times SL^2 + 0,065 \times SL$
<i>R</i> ²	90,7%	84,2%	83,5%	86,9%
<i>R</i> _a ²	90,6%	84%	83,3%	86,7%
<i>CV</i>	16,8%	16,2%	26,6%	18,9%

Legend: *CV* (Coefficient of Variation), *M* (Mean), MB (Mega Byte), *p* (Probability), *R*² (Coefficient of Determination), *R*_a² (Adjusted Coefficient of Determination), η^2 (Effect Size), *RRD* (Registration Request Delay), *SD* (Standard Deviation), *SDD* (Session Disconnect Delay), *SL* (Signalling Load), *SRD* (Session Registration Delay), *QoE* (Quality of Experience).

Table 2 Summary of results for impacts of signalling load on SIP call control signalling performances and *QoE* when the algorithm is activated

Signalling load (MB)	Overall user <i>QoE</i> (<i>M</i> (<i>SD</i>))	Human perception of <i>RRD</i> (<i>M</i> (<i>SD</i>))	Human perception of <i>SRD</i> (<i>M</i> (<i>SD</i>))	Human perception of <i>SDD</i> (<i>M</i> (<i>SD</i>))
0,00	5,00 (0)	5,00 (0)	5,00 (0)	5,00 (0)
0,73	5,00 (0)	5,00 (0)	5,00 (0)	5,00 (0)
1,47	5,00 (0)	5,00 (0)	5,00 (0)	5,00 (0)
2,20	4,97 (0,183)	5,00 (0)	4,93 (0,254)	5,00 (0)
2,93	4,77 (0,430)	5,00 (0)	4,07 (0,583)	4,87 (0,346)
3,66	4,47 (0,681)	5,00 (0)	3,80 (0,664)	4,77 (0,430)
4,40	3,83 (0,648)	4,10 (0,548)	3,33 (0,661)	4,13 (0,507)
5,13	3,10 (0,662)	3,17 (0,592)	2,93 (0,691)	3,93 (0,583)
5,86	2,90 (0,712)	3,00 (0,695)	2,73 (0,740)	3,60 (0,724)
<i>F</i> (8, 270)	95,052	169,199	102,814	56,785
<i>p</i>	<0,001	<0,001	<0,001	<0,001
η^2	0,744	0,838	0,759	0,635
Regression model	$4,976 - 0,098 \times SL^2 + 0,188 \times SL$	$4,871 - 0,124 \times SL^2 + 0,381 \times SL$	$4,951 + 0,447 \times SL^3 - 0,324 \times SL^2 + 0,031 \times SL$	$4,972 - 0,067 \times SL^2 + 0,147 \times SL$
<i>R</i> ²	72,7%	77,8%	74,7%	61,6%
<i>R</i> _a ²	72,5%	77,7%	74,4%	61,3%
<i>CV</i>	11,3%	9,0%	12,5%	8,8%

Legend: *CV* (Coefficient of Variation), *M* (Mean), MB (Mega Byte), *p* (Probability), *R*² (Coefficient of Determination), *R*_a² (Adjusted Coefficient of Determination), η^2 (Effect Size), *RRD* (Registration Request Delay), *SD* (Standard Deviation), *SDD* (Session Disconnect Delay), *SL* (Signalling Load), *SRD* (Session Registration Delay), *QoE* (Quality of Experience).

Table 3 Two-way ANOVA results indicating the influence of proposed algorithm on SIP signalling performances and *QoE*

	Overall user <i>QoE</i>	User perception of <i>RRD</i>	User perception of <i>SRD</i>	User perception of <i>SDD</i>
LOAD*ALGORITHM				
<i>F</i> (8, 540)	65,346	33,812	55,101	247,544
<i>p</i>	<0,001	<0,001	<0,001	<0,001
η^2	0,500	0,341	0,458	0,791
LOAD				
<i>F</i> (8, 540)	402,641	349,517	405,052	768,026
<i>p</i>	<0,001	<0,001	<0,001	<0,001
η^2	0,861	0,843	0,861	0,922
ALGORITHM				
<i>F</i> (1, 540)	872,001	441,961	901,888	1293,394
<i>p</i>	<0,001	<0,001	<0,001	<0,001
η^2	0,626	0,458	0,633	0,712

Legend: *CV* (Coefficient of Variation), *p* (Probability), η^2 (Effect Size), *RRD* (Registration Request Delay), *SDD* (Session Disconnect Delay), *SRD* (Session Registration Delay), *QoE* (Quality of Experience).

The percent of variability of *QoE* can be explained by knowing signalling load by approximately 72,7%, whereas the calculated *CV* is approximately 11,3%. Further on, the percent of variation of user perception of *RRD*, *SRD* and *SDD* can be accounted for by the knowledge of call control signalling load by approximately 77,7%, 74,4% and 61,3%, while the calculated *CV* is approximately 9%, 12,5% and 8,8%, respectively. According to the aforementioned results, it can be noticed that the models given in Tab. 1 represent quantification of mutual relationships of signalling load and human perceptions of SIP signalling performances and *QoE* for VoIP when the algorithm is activated.

Final aim of this paper was to quantify the relationships between human perception of SIP signalling performances and *QoE*, when the algorithm is used. In order to test hypothesis H4 and to investigate how each of these human perceptions is connected to *QoE*, multiple linear regression analysis has been conducted with three predictors being the considered user perceptions. The resulting linear model is given by Eq. (2). The percent of variability of global *QoE* can be explained by the knowing user perception of *SRD* and *SDD* by approximately 65,7%, whereas the obtained coefficient of variation is about 12,5%.

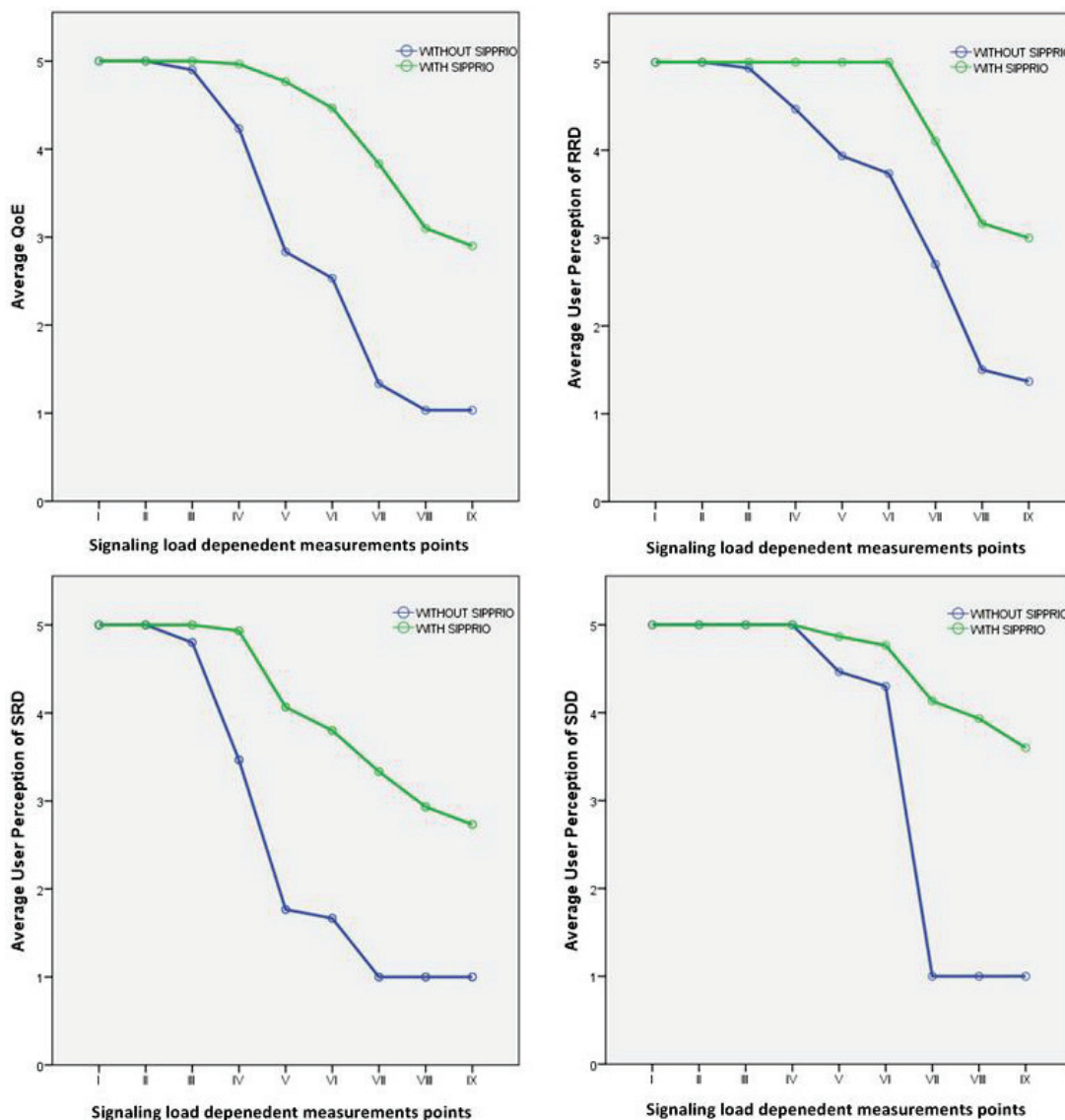


Figure 1 Graphical representation of dependence of human perception of *RRD*, *SRD* and *SDD*, and *QoE* on call control signalling load and algorithm implementation

$$QoE = 0,600 \cdot UPRRD + 0,398 \cdot UPSDD. \quad (2)$$

According to the aforementioned results, it can be concluded that the proposed model represents quantification of mutual relationships between the human perceptions of SIP signalling performances and *QoE* for VoIP when the algorithm is used, since, as indicated earlier, models characterized by high *R²* factor are outstanding. In addition, according to obtained model, we determined the importance of different dimensions in terms

of considered human perceptions and *QoE* in this scenario. Namely, human perception of *RRD* and *SDD* varies to the extent to which they affect the *QoE* (going from most to least important), while human perception of *SRD* has no influence of the *QoE*. This implies that assumed order of importance in hypothesis H4 is not supported.

This result may be explained by the fact that the human perception of signalling performance metrics is impacted by the proposed algorithm, which reduces the duration of call control signalling procedures in general. Assigning the

lowest priority to packets containing SIP message for establishing call control session blocks the session setup, which improves the overall *QoE* especially under overburden conditions. Since users do not agree with the service breakdown once they have begun a session, they rather accept the session to be refused whenever it is impossible to ensure appropriate service quality. Moreover, by assigning the highest priority to packets containing SIP message for terminating the call control session reduces the duration of call teardown procedure and its importance on overall *QoE*. Reducing the importance of human perception of *SRD* and *SDD* increases the importance of *RRD*. Therefore, the obtained linear regression model may be used by service providers for predicting the service availability.

5 CONCLUSION AND FUTURE WORK

This paper aimed to explore and gain more comprehensive understanding of the interaction between the SIP call control signalling, *QoE* for VoIP and human perceptions of SIP signalling performances. Additionally, the aim was to verify the algorithm for differentiated treatment of SIP messages from the user-centric aspect. In order to achieve this, we have conducted a survey of research studies in the area of influence of SIP signalling load on service quality and *QoE*. Further on, we have conducted the experimental study to obtain data for describing the relationship between SIP call control signalling, human perception of SIP signalling performances and *QoE* including both scenarios, the algorithm for differentiated treatment of SIP messages being activated and not. Results obtained from these experiments are used to explain the *QoE* and quality perception of register, call setup, and teardown procedure while using VoIP service under different signalling load conditions. Finally, we have conducted the statistical analysis of the obtained results using descriptive statistics, regression analysis, and one- and two-way ANOVA.

The results obtained in this paper imply that call control signalling has its share and responsibility in influencing user *QoE* with VoIP service. It is shown that call control signalling load influences the human perception of SIP signalling performances and *QoE*. Thus, the model that describes the relationship between human perception of SIP signalling performances and *QoE* with VoIP is proposed. Moreover, it is presented that the algorithm for differentiated treatment of SIP messages influences human perception of SIP signalling performances and *QoE* with VoIP, in terms that it decreases the strong influence of signalling load on analysed variables. Thus, another model for quantifying mutual relationships of analysed human perceptions, and *QoE* when the proposed algorithm is activated in order to let us to recognize the value of distinct SIP signalling performances. The knowledge and results obtained in this study may be utilized by the service providers to improve their charging schemes (e.g., decide whether or not to charge the call setup), and further improve user experience (e.g., predict and determine acceptable thresholds for service availability) given the fact that we live in a world where human perceptions, expectations and experiences carry a greater value.

Since this is a contemporary and attractive area of research, a number of open issues may be identified. One

of them chosen to be addressed in our future work is directed to the investigation of service differentiation influence on *QoE*. It may be achieved by using the algorithm for differentiated treatment of SIP messages. Its influence on *QoE* for native and Web-based unified communication services may be investigated within our future research.

6 REFERENCES

- [1] Barakovic, S. & Skorin-Kapov, L. (2015). Multidimensional Modelling of Quality of Experience for Mobile Web Browsing. *Computers in Human Behavior*, 50, 314-332. <https://doi.org/10.1016/j.chb.2015.03.071>
- [2] Batteram, H., Damm, G., Mukhopadhyay, A., Philippart, L., Odysseos, R., & Urrutia-Valdes, C. (2010). Delivering Quality of Experience in Multimedia Networks. *Bell Labs Technical Journal*, 15, 175-194. <https://doi.org/10.1002/bltj.20431>
- [3] Barakovic, J., Bajric, H., Kos, M., Barakovic, S., & Husic, A. (2011). Prioritizing signaling information transmission in next generation network. *Journal of Computer Networks and Communications*, 011. <https://doi.org/10.1155/2011/470264>
- [4] Barakovic Husic, J., Bajric, H. & Barakovic, S. (2012). Evolution of Signalling Information Transmission. *ISRN Communications and Networking*, 2012. <https://doi.org/10.5402/2012/705910>
- [5] Barakovic Husic, J., Perenda, E., Hadzialic, M., & Barakovic, S. (2013). Performance modelling and optimization of IP Multimedia Subsystem. *Proceedings of the UKSim AMSS 7th European Modelling Symposium on Mathematical Modelling and Computer Simulation (EMS)*, Manchester, UK. <https://doi.org/10.1109/EMS.2013.103>
- [6] Barakovic Husic, J., Hidic, A., Hadzialic, M., & Barakovic, S. (2014). Simulation-based Optimization of Signaling Procedures in IP Multimedia Subsystem. *Proceedings of the 15th Conference of Open Innovations Association FRUCT*, St. Petersburg, Russia. <https://doi.org/10.1109/FRUCT.2014.6872433>
- [7] Barakovic Husic, J., Barakovic, S., & Kasumovic, S. (2018). Studying the Impact of SIP Message Differentiation on Quality of VoIP Session Control Procedures. Manuscript submitted for publication.
- [8] Barakovic Husic, J., Barakovic, S., Kovac, S., Delija, A., & Kovac, E. (2017). An Investigation of the Influence of SIP Message Differentiation on the Quality of Unified Communications. *Proceedings of the 26th International Conference on Information, Communication and Automation Technologies (ICAT)*, Sarajevo, B&H. <https://doi.org/10.1109/ICAT.2017.8171639>
- [9] Voznak, M. & Rozhon, J. (2013). Approach to stress test in SIP environment based on marginal analysis. *Telecommunication Systems*, 52, 1583-1593. <https://doi.org/10.1007/s11235-011-9525-1>
- [10] Goncalves, D., Amaral, A., Costa, A., & Sousa, P. (2016). Towards automated test and validation of SIP solutions. *Telecommunication Systems*, 61, 579-590. <https://doi.org/10.1007/s11235-015-0015-8>
- [11] Alshamrani, M., Cruickshank, H., Sun, Z., Ansa, G., & Alshahwan, F. (2016). SIP Signaling Implementation and Performance Enhancement over MANET: A Survey. *International Journal of Advanced Computer Science and Applications*, 9, 191-202. <https://doi.org/10.14569/IJACSA.2016.070529>
- [12] Ali, A., Alshamrani, M., Kuwadekar, A., & Al-Begain, K. (2015). Evaluating SIP Signaling Performance for VoIP over LTE Based Mission-Critical Communication Systems. *Proceedings of the 9th International Conference on Next*

- Generation Mobile Applications, Services and Technologies*, Cambridge, UK.
<https://doi.org/10.1109/NGMAST.2015.68>
- [13] Elnashar, A., El-Saidny, M. A., & Mahmud, M. (2017). Practical Performance Analyses of Circuit-Switched Fallback and Voice over LTE. *IEEE Transactions on Vehicular Technology*, 66, 1748-1759.
<https://doi.org/10.1109/TVT.2016.2560962>
- [14] Kulikov, N. (2016). Analysis of Quality of Service Metrics in IMS Networks. *Automatic Control and Computer Sciences*, 50, 37-45.
<https://doi.org/10.3103/s0146411616010077>
- [15] Montazerolghaem, A., Yaghmaee, M. H., & Leon-Garcia, A. (2017). OpenSIP: Toward Software-Defined SIP Networking. *IEEE Transactions on Network and Service Management*, PP. <https://doi.org/10.1109/TNSM.2017.2741258>
- [16] Kolhar, M., Alameen, A., & Gulam, M. (2017). Performance evaluation of framework of VoIP/SIP server under virtualization environment along with the most common security threats. *Neural Computing and Applications*, 1-9.
<https://doi.org/10.1007/s00521-017-2886-y>
- [17] Nguyen, D. T., Nguyen, K. K., Khazri, S., & Cheriet, M. (2016). Real-time optimized NFV architecture for interworking WebRTC and IMS. *Proceedings of the 17th International Telecommunications Network Strategy and Planning Symposium (Networks)*, Montreal, QC, Canada.
<https://doi.org/10.1109/NETWKS.2016.7751157>
- [18] Malas, D. & Mortan, A. (2011). Basic Telephony SIP End-to-End Performance Metrics (Tech. Rep. RFC 6076). Retrieved from <https://tools.ietf.org/html/rfc6076>
- [19] Fathi, H., Chakraborty, S. S., & Prasad, R. (2009). *Voice over IP in Wireless Heterogeneous Networks: Signalling, Mobility and Security*. Springer Netherlands.
- [20] Agbinya, J. I. (2010). *IP Communications and Services for NGN*. Boca Raton: CRC Press.
<https://doi.org/10.1201/9781420070910>
- [21] Happenhofer, M., Egger, C., & Reichl, P. (2010). Quality of Signaling: A New Concept for Evaluating the Performance of non-INVITE SIP Transactions. *Proceedings of the 22nd International Teletraffic Congress (ITC)*, Amsterdam, Netherlands. <https://doi.org/10.1109/ITC.2010.5608733>
- [22] Brajdic, A., Suznjevic, M., & Matijasevic, M. (2009). Measurement of SIP Signaling Performance for Advanced Multimedia Services. *Proceedings of the 10th International Conference on Telecommunications (ConTEL)*, Zagreb, Croatia.
- [23] Tang, J., Davids, C., & Cheng, Y. (2008). A Study of an Open Source IP Multimedia Subsystem Testbed. *Proceedings of the 5th International ICST Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine)*, Hong Kong, China.
<https://doi.org/10.4108/ICST.QSHINE2008.3952>
- [24] Batalla, J. M., Sliwinski, J., Tarasiuk, H., & Burakowski, W. (2010). Impact of Signaling System Performance on QoE in Next Generation Networks. *Journal of Telecommunications and Information Technology*, 117-130.
<https://doi.org/10.1007/s12243-008-0060-3>
- [25] Abhaywardhana, V. S. & Babbage, R. (2007). A Traffic Model for the IP Multimedia Subsystem (IMS). *Proceedings of the 65th Vehicular Technology Conference (VTC Spring)*, Dublin, Ireland. <https://doi.org/10.1109/VETECS.2007.171>

Contact information:

Jasmina BARAKOVIĆ HUSIĆ, PhD, Associate Professor
 (Corresponding author)
 Faculty of Electrical Engineering, University of Sarajevo
 Zmaja od Bosne bb,
 71000 Sarajevo, Bosnia and Herzegovina
jasmina.barakovic@etf.unsa.ba

Sabina BARAKOVIĆ, PhD, Assistant Professor
 Faculty of Traffic and Communications, University of Sarajevo
 Zmaja od Bosne 8,
 71000 Sarajevo, Bosnia and Herzegovina
barakovic.sabina@gmail.com

Seudin KASUMOVIĆ, MSc
 Bicom Systems
 Druge Tuzlanske brigade 13,
 75000 Tuzla, Bosnia and Herzegovina
seudin.kasumovic@gmail.com