Computer Science and Information Systems 12(2):587–605 DOI: 10.2298/CSIS140810018S

# Statistical User Behavior Detection and QoE Evaluation for Thin Client Services

Mirko Suznjevic[1], Lea Skorin-Kapov[1], and Iztok Humar[2]

[1] University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia,
{mirko.suznjevic, lea.skorin-kapov}@fer.hr,
[2] University of Ljubljana, Faculty of Electrical Engineering, Tržaška 25, 1000 Ljubljana, Slovenia
iztok.humar@fe.uni-lj.si,

**Abstract.** Remote desktop connection (RDC) services offer clients the ability to access remote content and services, commonly in the context of accessing their working environment. With the advent of cloud-based services, an example use case is that of delivering virtual PCs to users in WAN environments. In this paper, we aim to detect and analyze common user behavior when accessing RDC services, and use this as input for making Quality of Experience (QoE) estimations and subsequently providing input for effective QoE management mechanisms. We first identify different behavioral categories, and conduct traffic analysis to determine a feature set to be used for classification purposes. We propose a machine learning approach to be used for classifying behavior, and use this approach to classify a large number of real-world RDCs. We further conduct QoE evaluation studies to determine the relationship between different network conditions and subjective end user QoE for all identified behavioral categories. Results show an exponential relationship between QoE and delay and loss degradations, and a logarithmic relationship between QoE and bandwidth limitations. Obtained results may be applied in the context of network resource planning, as well as in making QoE-driven resource allocation decisions.
**Keywords:** user behavior, remote desktop connection, traffic classification, machine learning, QoE.

## 1. Introduction

With the advent of cloud computing and data centers offering virtual desktop solutions to end users, thin-client solutions are becoming an increasingly popular mechanism for accessing and interacting with remote content and services in an easy-to-maintain and cost effective manner [5]. In such environments, thin clients can be characterized as simple user interaction terminals, with the requirements for computational power, storage, and maintenance moved to the cloud [19]. Additional benefits include a reduction in the number of underutilized desktops hence leading to energy efficiency, wider access to data regardless of end user devices, and improved data security [2]. Today numerous solutions exist on the market supporting remote desktop connections (RDC), including those provided by Microsoft, Oracle, and Citrix.

A key challenge in making such solutions viable from an end user point of view is meeting the stringent network performance requirements dictating low delays and high response times [14]. Even small increases in delay may dramatically impact user perceived quality, an issue that is of particular concern in wireless and mobile networks [21]. As opposed to a standard, "local" desktop, whereby user inputs are locally processed and rendered nearly immediately, RDCs require inputs to be transmitted to a remote computer, processed, and returned to the thin-client [4]. Consequently, the screen updates and response times become a critical issue, in particular in WAN environments.

RDC traffic generally corresponds to encrypted video bitmaps exchanged between a remote virtual PC and a thin-client using a single server port, imposing challenges in terms of detecting end user tasks and applications that are being remotely run (e.g., a user editing a document, browsing the Web, or viewing audio/video content) [6]. RD protocols (e.g., Microsoft Remote Desktop Protocol, RDP) commonly run over TCP and offer a reliable connection. Different user behaviors while using RDC (in terms of conducted task or application being used) result in different traffic characteristics and different impacts of network performance on user perceived quality [21]. Therefore, a prerequisite in effectively managing the end user Quality of Experience (QoE) is related to determining the user behavior in the context of RDC for the purpose of accurately mapping between QoE and performance indicators (i.e., delay, bandwidth). Calyam et al. [2] stress the need to consider QoE-driven utility functions in the context of optimally allocating resources (i.e., CPU, memory, bandwidth) to virtual desktop flows.

In this paper, we focus first on detecting user behavior while using RDC based on employing statistical traffic analysis and machine learning. We compare achieved results with those reported in related work [6]. Further, going beyond related work, we apply our proposed approach to real-world RDC traffic traces collected at the Faculty of Electrical Engineering University of Ljubljana to study actual user behavior. It should be noted that this work is an extension of our paper published in [22] in which we presented our initial results regarding the algorithm for classification of application level behavior. This paper is extended in comparison with the previous one in several aspects. In addition to minor adjustments in the area of traffic analysis performed during the classification, the main novelty of this paper includes additional QoE studies based on the classification of application level behavior. In order to exploit knowledge regarding such traffic characteristics, we report on a series of QoE evaluation studies focused on relating user perceived quality with heterogeneous network conditions for different addressed user behavior categories. The studies have been done at two different locations and with different testing groups so as to ensure that results and observed trends are not dependent of the group under test. We have also presented the methodology for performing such tests to enable possible cross-validation with future studies. Analyses presented focus on the impact of network delay, packet loss and available bandwidth on QoE. We further perform regression modeling to obtain the models which capture the dependency of QoE of each of the single network parameters in a specific application context (i.e., activity performed in the application). We believe that these extensions, combined with our previous results, present a firm foundation for QoE-based network management algorithms which do not take into account only network parameters, but also application level context. Hence, combined with the potential to identify user behavior in the network, the results of the QoE studies provide the opportunity to make effective QoE estimations for different behavior categories under various network conditions. Such QoE estimations may then be further used for management mechanisms such as resource planning or dynamic resource (re)allocation in order to assure high QoE and achieve a satisfied customer base [13, 25, 18].

The paper is structured as follows. In section 2 we give an overview of traffic classification approaches for thin client services, and further discuss studies addressing QoE in the context of such services. Section 3 focuses on traffic feature extraction and a traffic classification algorithm for RDC employing a machine learning approach. In section 4, we use our algorithm to analyze actual user behavior captured in 18.5 GB of real RDC traces in an academic setting. Section 5 reports on QoE studies addressing the relationship between network conditions and QoE for the chosen set of previously addressed user RDC behavior categories. Section 6 presents concluding remarks and future research.

## 2. Related Work

**Traffic classification**

While approaches such as those based on Deep Packet Inspection have been commonly used to identify and classify network traffic [20], a key drawback lies in the fact that traffic may be encrypted, preventing payload analysis. Additional drawbacks include necessary knowledge regarding payload formats, and potential government imposed privacy regulations. Approaches based on statistical traffic analysis present a different approach, whereby relevant traffic features (e.g., packet length and inter-arrival times) are extracted and analyzed to identify a particular application. Emmert et al. [7] previously reported on traffic patterns for different types of thin client users when working with popular office applications such as Microsoft Word, Excel, or PowerPoint. A thorough workload characterization and modeling (session arrival process, inter-arrival times, duration) of RDC traffic was done by Humar *et al.* [10]. The authors further confirmed self-similar characteristics of RDC traffic at packet level in [11].

In the past, Machine Learning (ML) algorithms have also been used for traffic classification purposes [15, 16]. Dusi *et al.* [6] propose a novel approach to RD application identification based on IP-level statistical traffic analysis and machine learning techniques, their goal being to make QoE estimations targeted towards verifying that Service Level Agreements are met. In their tests, the authors ran several applications over an RDP connection, categorized as audio (e.g., VLC, WMP, Skype), video (e.g., Flash video, WMP), and data (e.g., Web browsing, Adobe reader). The authors extracted traffic features (bit rate and packet rate at IP and TCP levels, TCP payload length, and number of observed packets) and evaluated different statistical classification techniques in terms of accuracy. Their results show that on-the-fly application identification for thin-client flows can achieve over 90 percent accuracy (when applying Support Vector Machines and Decision Tree algorithms, and considering a window size of 10 s) even when the testing set includes applications for which their classifier was not trained.

**Relation to QoE studies**

The need for QoE models and the general challenges in providing effective QoE management for cloud-based services such as RDC are highlighted in [9]. Consequently, previous research has studied RDC traffic and its relation to user perceived QoE. When considering Desktop as a Service (Daas), Deboosere *et al.*[5] highlight two important aspects impacting user experience as being high performance of the thin client protocol hence resulting in crisp interactivity and fluent audiovisual output, and sufficient server-side resource allocation.

The user experience regarding RDC was quantified by Tolia *et al.* [23] through operation response times on different bandwidths (10 and 100 Mbit/s) and for different types of office applications (typing, tracking changes, creating slides for presentation and manipulating photos). Staehle *et al.* [21] studied how different network parameters (packet loss, jitter and delay) influence subjective and objective QoE in a controlled Citrix-based testbed environment, whereby they evaluate task completion time for different applications (using MS Word, Excel, and PowerPoint) and relate this metric to user satisfaction. In their subsequent work, the authors again focused on Citrix-based thin client architectures, widely used in professional environments, and addressed how to optimize QoE (in terms of task completion time) using different thin client settings [19]. Objective assessment of QoE for thin client based virtual desktop clouds is also addressed in recent work by Xu *et al.* [26].

Further, while Dusi *et al.* proposed a mapping between QoE scores and round trip time (RTT) thresholds for different types of RD applications, a more extensive study addressing the QoE of remote desktop users was conducted by Casas *et al.* [4]. In a laboratory study involving 52 participants, the authors looked to identify relationships between network performance parameters (i.e., delay, bandwidth) and end user subjective quality scores for four typical remote desktop tasks found in enterprise scenarios and using a Citrix-based virtual remote desktop setup: text typing (e.g., email and document editing), screen scrolling (e.g., web browsing), drag and drop of images, and menu browsing (e.g., menu selection). The authors further provide a traffic characterization whereby they constitute that the four aforementioned tasks present nearly identical traffic patterns in the uplink, while in the downlink typing sends much less data than the other three tasks, which share a common packet size behavior. Differences in throughput are linked with different requirements with regards to sending screen updates. Given that RDC traffic uses TCP as a transport protocol, network delay has been found to have the most significant impact on end user QoE. The authors' findings show that, in general, RTT should be kept below 150 ms to assure good QoE. Furthermore, the authors compared the behavior of remote desktop users confronted with variable network conditions to the behavior of users on standards local desktops and found that the the duration to complete a certain task using remote desktop can take up to 3 times longer (given round trip times up to 500 ms) than on a local desktop.

The exploitation of application identification based on statistical analysis is addressed by Arumaithurai *et al.* [1] in the context of QoE-aware scheduling. The authors propose a solution for providing preferential treatment to certain thin-client flows over other thin-client flows traversing the same intermediate network node based on the QoE-driven resource requirements of the application being run by a user at a given time.

In the following sections, we first address user behavior detection while using RDC services, and then further link these different behavior categories to QoE studies to identify network performance requirements from an end user perspective.

# 3. User behavior detection based on machine learning

With regards to user behavior prediction for RDC, we build on the work reported by Dusi *et al.* [6]. Given that their results showed that the Decision Tree ML algorithm (tested using the freely available WEKA software [24]) provided very good results in terms of traffic classification accuracy, we have opted to use this algorithm in our study.

## 3.1. Methodology

The first step in our work was to train an ML algorithm to classify different types of RDC behavioral activities, to be subsequently used for the analysis of real RDC traffic traces. The behavior labeled traces which were used for training the Decision Tree algorithm were collected at the Department of Telecommunications, Faculty of Electrical Engineering and Computing, University of Zagreb. The capture was performed on a Dell OptiPlex 390 computer (configuration: i3@3,3 GHz, 4GB RAM, ATI Radeon HD 6450) connected to another Dell OptiPlex 390 via a 100 Mbit LAN. The behavior categories we have addressed are as follows:

- idle - no actions for 10 seconds, static screen (desktop image);
- document editing - editing of a word document (including text writing, picture pasting, text copying etc.);
- browsing - searching for accommodation on the www.booking.com web page;
- audio - listening to a 128 bit/s online radio station from www.radio365.com;
- video - watching a 10 minute full-screen movie on www.youtube.com.

The measurements for each of the behavior categories lasted 10 minutes, i.e., during those 10 minutes, only a given action was being performed. The actions recorded were performed by student volunteers of the Faculty of Electrical Engineering and Computing. We note that previous cited efforts [6, 4] have not considered the category of *idle*, even though this is commonly observed in RDC traffic. The reason is that previous efforts focused either on identifying certain applications being run over an RDC or studying end user QoE, while our goal is to study overall user behavior exhibited when using RDCs. Following identification of a number of traffic features to be used for classification, we conducted a traffic analysis and specified a decision tree algorithm. We then used our collected traces to train the ML algorithm. Finally, we validated the algorithm using a validation dataset.

## 3.2. Traffic feature extraction

Features are attributes of flows calculated over multiple packets, used to train a ML classifier in associating sets of features with given application types [15]. Behavior labeled traces were processed such that for an epoch (time window) of 10 seconds, six features were extracted:

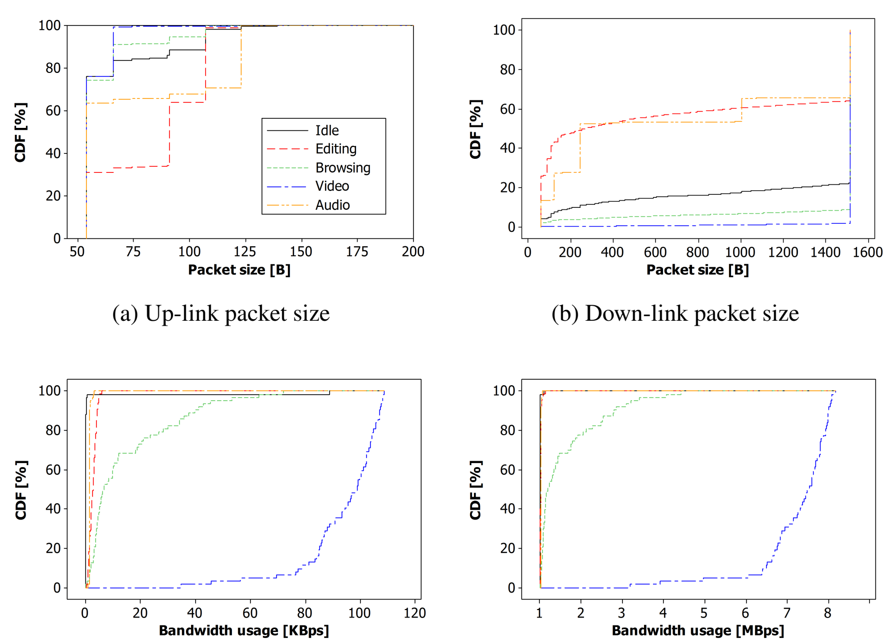
- up-link packet number;
- up-link average packet size;
- up-link average bandwidth usage;
- down-link packet number;

– down-link average packet size;
– down-link average bandwidth usage;

By *up-link* we refer to packets originating from the RDC client, while *down-link* refers to packets originating from the RDC server. The reason for choosing an epoch of 10 seconds is due to the fact that previous research has shown this approach as providing good results in terms of accuracy [6]. Traffic features were extracted from the network traffic traces using a custom built Java-based parser.

### 3.3. Machine learning approach

Following traffic capture, we analyzed the traces to determine different traffic patterns for the identified behavioral categories. With regards to the traffic characteristics, we note that the most distinguishing feature was found to be the number of down-link packets. Figure 1 portrays CDFs for packet size in both uplink and downlink directions, as well as CDFs for bandwidth usage across the different behavioral categories. In the case of traffic corresponding to the client being idle, only keep-alive packets were sent from server to client. As a result, we subsequently classified all traffic with <5 packet/s as idle. Video clearly represented the most resource demanding traffic, with average bit rates over 7 Mbytes/s (i.e., over 56 Mbit/s). We once again note that we used full-screen video. Such high re-




(a) Up-link packet size (b) Down-link packet size

(c) Up-link bandwidth usage (d) Down-link bandwidth usage

**Fig. 1.** Traffic analysis of RDC traces

source demands resulted in poor video quality in terms of image freezes and jerkiness, linked also to the fact that TCP usage is not optimized for real-time media. Browsing and editing exhibited highly variable traffic characteristics, with browsing resulting in higher down-link bandwidth utilization. Nevertheless, we presume that this is highly dependent on the end user interaction behavior resulting in screen updates. With regards to audio, we used a 128-bit radio channel and a static screen, i.e., there was no need for screen updates. We further note that all behavioral categories exhibit similar up-link characteristics (mostly composed of ACK packets), as also reported by Casas et al [4]. In the case of video, the higher bandwidth usage is due to the large number of ACKs being transmitted.

Having statistically analyzed captured traffic, we used the freely available WEKA Java library (developed at the University of Waikato, New Zealand) [24]. While the WEKA software contains a large collection of tools for data processing, machine learning, and data analysis, we focused on support for implementation of the Decision Tree algorithm (noted in WEKA as J.48). The algorithm we specified in our study is portrayed in Figure 2, relying on the identified feature set for classification purposes. During the training phase, the algorithm uses extracted features from the traffic samples to build the mapping functions from the samples to the behavior category [1]. We used the approach proposed in

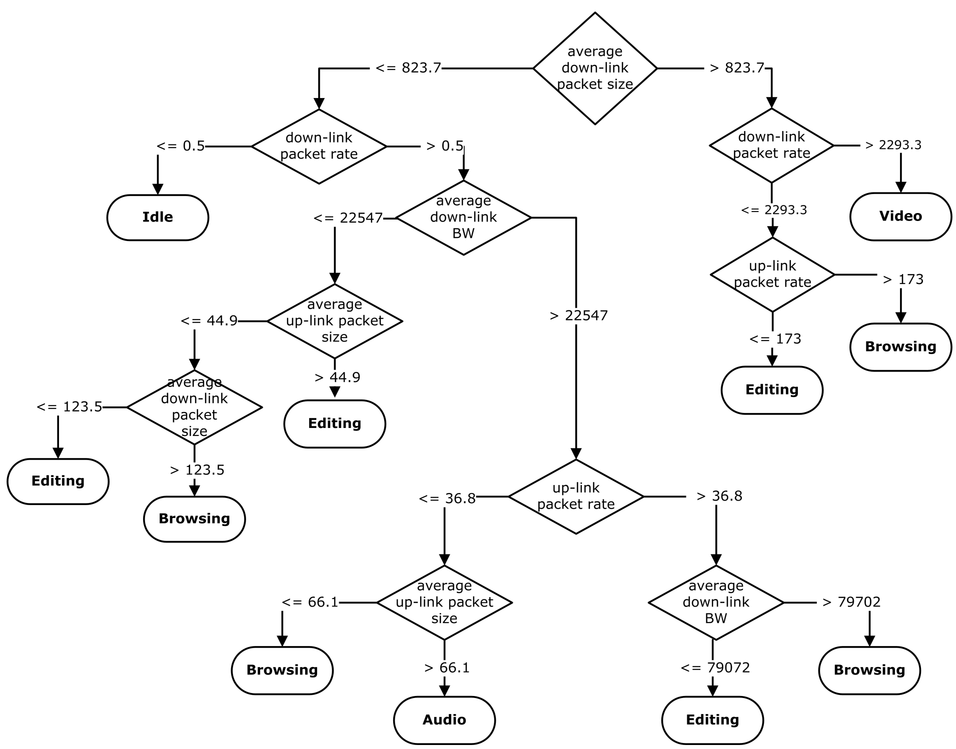



**Fig. 2.** A view of the Decision Tree algorithm used for traffic classification

[1] Interested readers may contact the authors for access to the data samples used for training purposes.

[6] to split the samples into fixed time epochs (i.e., time windows) of 10 s and extract the features for each epoch. However, while in [6] the authors use only down-link traffic for algorithm training, we incorporate also up-link samples.

Having completed the training phase, we generated an annotated dataset with the addressed behavioral categories for validation purposes. The accuracy in terms of correct classification (in terms of epochs) was found to be 78.0%, while the accuracy in terms of correctly classified bytes was found to be 89.7%. While this is somewhat lower than the accuracy reported by [6], we note that this may be attributable to the fact that we consider document editing and web browsing as different categories, whereas the cited authors considered a general "data" category (including web browsing, Adobe Reader, and Microsoft PowerPoint). Furthermore, the authors do not consider *idle* periods in their traffic classification. Finally, while the authors considered only down-link traffic, they used an extended feature set for classification purposes as compared to our set.

## 4. Analysis of real RDC traces

Following training and validation of the machine learning algorithm, we apply the algorithm to real RDC traces to study actual user behavior in a real world scenario (outside of a laboratory environment). To obtain empirical traffic traces, we collected packet-level traces on a 100 Mbit/s Ethernet link that connects the Faculty of Electrical Engineering, University of Ljubljana (FE) to the external Internet, as shown in Figure 3. The traces

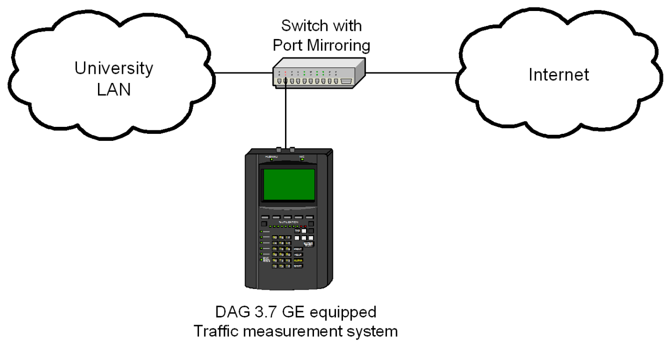



**Fig. 3.** Testbed environment for collecting RDC traces at FE, Ljubljana

were collected from Sept. $24^{th}$-Nov. $16^{th}$ 2012. We used a traffic capturing system on a personal computer running Linux OS and with a DAG network interface card. The traces were collected, filtered and stored in .cap format. We focused on the most frequently used RDC traffic, Microsoft Remote Desktop Protocol, which was recognized by transport layer (TCP) port numbers, configured to 3389 by default. To manage long traces, the individual files were limited to 0.5 GB. The traces contain RDC flows generated by students and staff of the Laboratory of Telecommunications, FE. The measurement procedure resulted in a total of 37 traffic traces which comprised 18.5 GB of RDC traffic. Each trace consisted of multiple RDCs and corresponding flows. A total of 1364 sessions were established when summarized across all traces (excluding sessions with a session length less than 10 s).

Figure 4 shows the CDF for the different session lengths that were collected in the traces. Results show that while the majority of sessions were of short length, there is a significant portion of longer sessions, some lasting up to multiple days. This is indicative of the fact that users established RDCs and left them open for extended periods of time, even though the connections were for the most part idle.

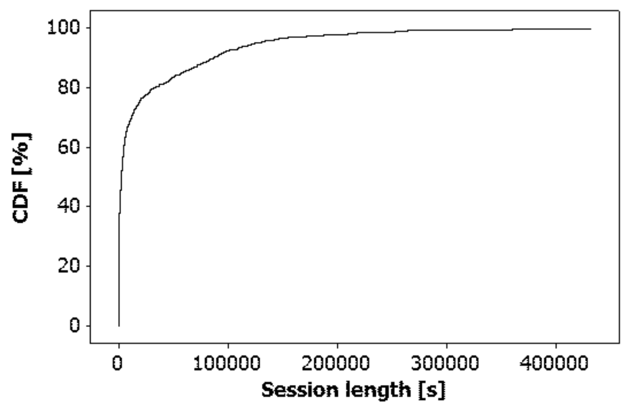



**Fig. 4.** CDF portraying session lengths for collected traces

Figure 5 summarizes the duration of total time that users spent engaged in the given behaviors across all considered traffic traces. Results clearly show that for the majority of time (92.91 %), the RDC was idle, indicating that no end user activity was detected. Document editing (6.14 %) proved to be the most common activity, with small percent-

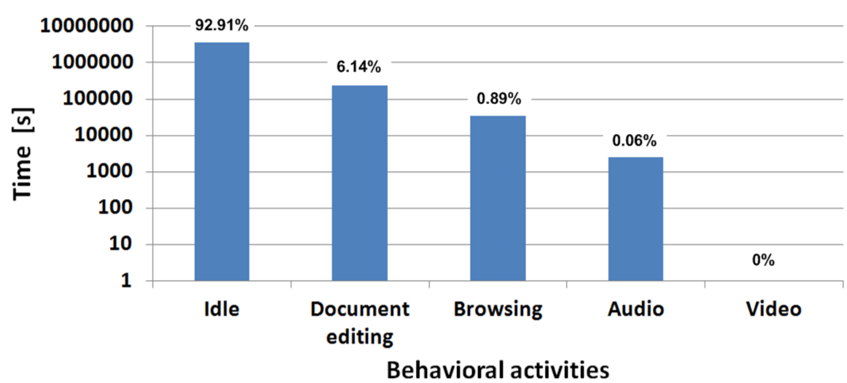



**Fig. 5.** Total time users spent engaged in a given behavioral category (logarithmic scale)

ages of time spent browsing (0.89 %) and listening to audio (0.06 %). Throughout the entire collection of traces, there was no detection of behavior indicating that end users were running video applications.

This supports the assumption that the users involved in this study generally did not watch video content via a remote desktop connection, most likely due to both the nature of the tasks that users perform when using RDC, and also due to poor video quality resulting from network delays. While delay is not a determining issue in LAN scenarios, the use of RDC services in WAN scenarios clearly leads to the increased impact of RTT.

## 5. QoE evaluation

Following the classification and analysis of different RDC behavioral categories in the previous sections, a key concern in aiming to meet end user quality expectations is an understanding of the impact of underlying network performance on perceived quality. In this section we therefore focus on quantifying the relationship between network performance and QoE, and compare our results to those obtained in related studies.

### 5.1. Methodology

In order to study the impact of different network conditions on end user QoE for different addressed behavioral categories, we have conducted a series of subjective tests. The tests were conducted across two different laboratories: QoE evaluations for document editing, web browsing, and streaming video were conducted at laboratory facilities at the Faculty of Electrical Engineering University of Ljubljana, while QoE evaluations of audio streaming were conducted at the University of Zagreb, Faculty of Electrical Engineering and Computing.

Both laboratory test environments were setup in the same way and involved two PCs connected via a third PC emulating various test conditions. The measurements were performed using MS Windows RDP, with one PC acting as a client from which the user establishes RDC connections and the other acting as the remote system. The workstations used were high performance PCs that did not influence the measuring processes with their characteristics, for their impacts were negligible in comparison to the network characteristics. In all cases, screen resolution was set to 1024 x 768 with 16 bit color depth. Default settings of compression were used while caching and data encryption were turned off.

Users taking part in the evaluation were asked to complete a set task under given test conditions, after which they provided overall QoE ratings using a five point Absolute Category Rating scale (1-bad, 2-poor, 3-fair, 4-good, and 5-excellent), commonly used as a de facto standard in QoE studies and specified in ITU-T Recommendation P.800.1 [12] (referred to as Mean Opinion Score, MOS). Users worked on each task individually, while a test administrator supervised the testing process and collected the results in a specified form. The users were not aware of the set parameters, and their evaluation was based on their own experiences with the same services when run locally on their PC. The following tasks were considered:

- document editing - users were asked to edit a word document (including text writing, picture pasting, text copying etc.);

- web browsing - users were asked to open a web browser and browse through a predefined set of web pages;
- audio - users listened to a song streamed for a duration of 20 seconds per test configuration, encoded using AAC-LC at an average bit rate of 256 kbit/s, sampling rate 44100 Hz, played by the GOM player;
- video - users watched a full-screen cartoon for a duration of a 60 seconds per test configuration.

While previous studies have shown that delay in terms of RTT is the main performance parameter to be considered, we conduct tests also under different bandwidth and loss values in order to extend the parameter consideration and provide a clearer view on the QoE impacts of different conditions. What is clear is that given that TCP is being used as the underlying transport protocol, packet loss will result in retransmissions and observed delays or retransmission time outs. Different test conditions were manipulated for the different behavioral categories in order to try and test all possible MOS scores, ranging from excellent to poor, and to find the behavior-specific lower bounds in terms of acceptable network degradations. The following test conditions were manipulated, ranging from a high-speed and low delay network to a slow network with high delay and packet loss (the rational being that such conditions may still found today during traffic peaks in mobile networks):

- Document editing: delay: 0 ms, 100 ms, 200 ms, 300 ms, 500 ms; loss: 1%, 2%, 5%, 10%; bandwidth: 9.6 kbit/s, 56 kbit/s, 128 kbit/s, 2 Mbit/s
- Web browsing: delay: 0 ms, 100 ms, 200 ms, 300 ms, 500 ms; loss: 1%, 2%, 5%, 10%; bandwidth: 56 kbit/s, 128 kbit/s, 2 Mbit/s, 10 Mbit/s
- Audio streaming: delay: 0 ms, 100 ms, 200 ms; loss: 1%, 2%, 5%; bandwidth: 128 kbit/s, 2 Mbit/s, 10 Mbit/s
- Video streaming: delay: 0 ms, 100 ms, 200 ms; loss: 1%, 2%, 5%; bandwidth: 128 kbit/s, 2 Mbit/s, 10 Mbit/s

We note that the indicated values were set symmetrically in both uplink and downlink directions, i.e. 100 ms delay corresponds in fact to 200 ms RTT. Further, the tests were designed so as to test the impact of each network performance parameter individually. Hence, in order to test the impact of a given parameter, the other parameters which were not being manipulated were set to optimal values (e.g., delay set to 0 ms and loss set to 0% while testing the impact of reduced bandwidth). In the case of testing delay and loss impact, no bandwidth limitations were imposed, hence users had access to a 100 Mbit/s LAN.

**Participants.** For the tests conducted in Ljubljana, participant demographics are as follows: $N = 20$, four females and 16 males with the average age of 22.1 years. All the participants were students and had a strong computer background and experience in working with the reference services, so they were able to evaluate perceived quality appropriately. For the tests conducted in Zagreb, demographics were as follows: $N = 20$, again four females and 16 males, with an average age of 29.4 years. Participants were employed as researchers at the Department of Telecommunications and all have a strong computer background.

### 5.2. Results and discussion

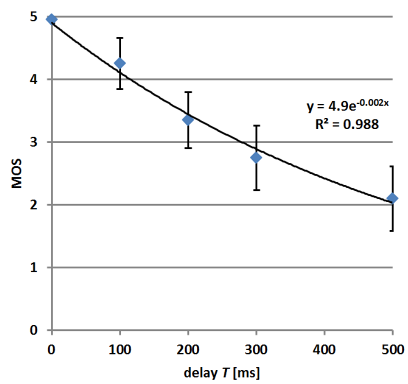
Figure 6 and Figure 7 portray the results of the described subjective tests for document editing and web browsing. Figure 8 depicts the results for audio and Figure 9 the results for video. As expected, document editing and web browsing proved to be activities during which users were much more tolerant towards network impairments than in the case of audio and video streaming. While in ideal network conditions document editing and web browsing showed "excellent" MOS scores, this was not the case for audio and video streaming. In particular, the quality of video streaming was rated between fair and poor for cases with no extra impairments imposed, indicating that the employed TCP-based RDC solution is not applicable in this case. In the case of audio, RTT delays of 200 ms already resulted in MOS falling below 2, while document editing and web browsing were more tolerant. We note that in the case of web browsing, the actual delay users perceive as regarding page loading is the sum of emulated delays and Internet access delays. Given that the local desktop was connected to the Internet via high-speed LAN access, we consider this additional delay to be negligible. The focus was also on scrolling through web pages and changing tabs, which are activities impacted by the RDC.

With respect to packet losses, document editing tolerated even up to 10% loss, web browsing up to 5% loss, and audio up to 1% loss in each direction (downlink and uplink). As regards bandwidth consumption, document editing was the least bandwidth-intensive application, while both web browsing and streaming audio resulted in MOS values above 3 at available bandwidths of 2 Mbit/s or greater.

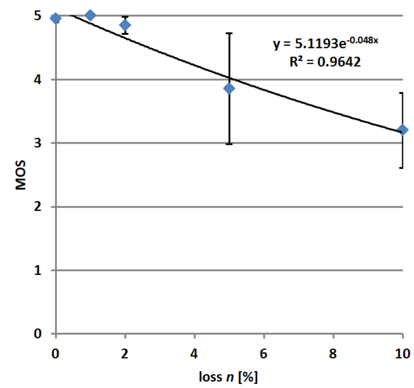
What is clear is that TCP is not efficient in network environments with limited bandwidth, high delay, and packet loss. Hence, newer solutions (e.g., newer versions of RDP) are focusing on incorporating support for User Datagram Protocol (UDP) as well, in particular for multimedia streaming services that are tolerant to a certain amount of packet loss, but intolerant to high delays and low throughput.

We further compare our findings with those mentioned previously in the context of related work. Casas and Schatz [3] conducted QoE studies for remote desktop tasks including texts typing, screen scrolling, drag and drop of images, and menu browsing, using a Citrix-based virtual remote desktop application. In our studies we additionally considered audio and video streaming, while relying on Windows RDP. Further, while the studies in [3] focus on the impact of RTT on QoE, we also manipulate loss and available bandwidth. Similarly to both [3] and [21], we found document (or text) editing to be the least delay-sensitive type of behavior, with acceptable MOS values (i.e., $\geq 3$) at up to 400 ms RTT, while Casas and Schatz found 350 ms RTT to be the limit for achieving MOS levels of 3 or above. Furthermore, while they found 250 RTT to be the limit for fair QoE (i.e. level 3) for page scrolling and menu browsing tasks, our results show MOS of approximately 3.5 at 200 ms RTT for a Web browsing task (which included also scrolling and menu browsing).

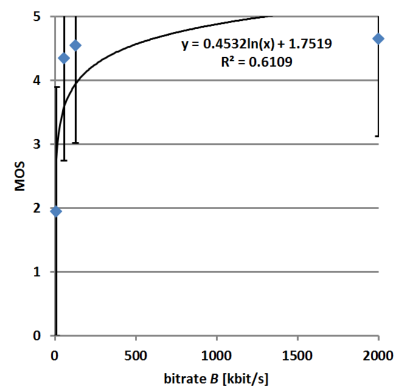
Dusi *et al.* [6] report only very generic results of their QoE study, where they found RTT between 100 and 400 ms as resulting in "sufficient" QoE (i.e., level 2 or 3 on a five point MOS scale) for web browsing (also in line with our findings), RTT between 120 and 450 ms as as resulting in sufficient QoE for audio (our results show slightly less tolerance in the case of audio streaming), and RTT between 50 and 70 ms resulting in sufficient QoE for video (also similar to our findings).




(a) MOS vs. delay



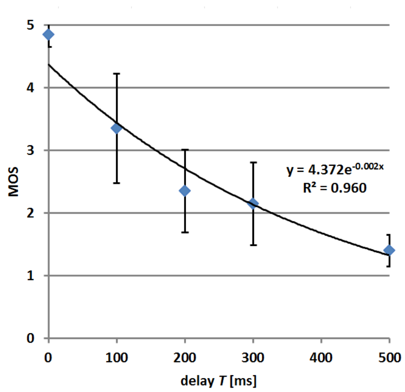

(b) MOS vs. loss



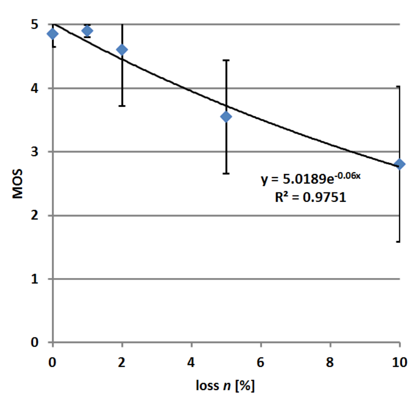

(c) MOS vs. available bandwidth

**Fig. 6.** MOS scores for document editing shown in relation to (a) delay, (b) loss, and (c) available bandwidth

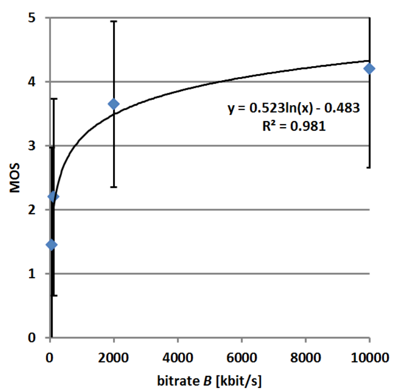
The graphs in Figures 6 through 9 portray curve fittings with a high coefficient of determination. A lower coefficient of determination is observed in the case of mapping MOS to available bandwidth for text editing. In all cases, we observe an exponential fitting in the cases of MOS mapped to delay and loss values. Such observations may be related to the generic IQX hypothesis postulated in [8], showing an exponential relationship between QoS and QoE, and stating that the change in QoE with respect to QoS degradation depends on the current QoE level. On the other hand, MOS values in relation to available bandwidth portray logarithmic relationships, also observed in previous QoE studies in the context of the Weber-Fechner Law (WFL), which studies the perceptive abilities of the human sensory system to a physical stimulus in a quantitative fashion, and states that the just noticeable difference between two levels of a certain stimulus is proportional to the magnitude of the stimuli [17].




(a) MOS vs. delay




(b) MOS vs. loss



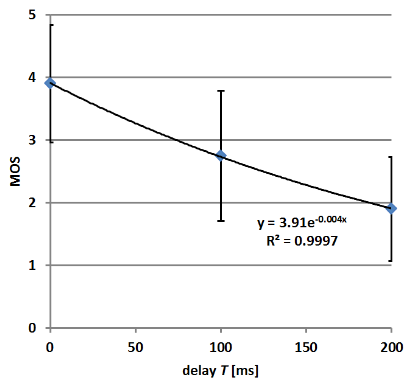

(c) MOS vs. available bandwidth

**Fig. 7.** MOS scores for web browsing shown in relation to (a) delay, (b) loss, and (c) available bandwidth

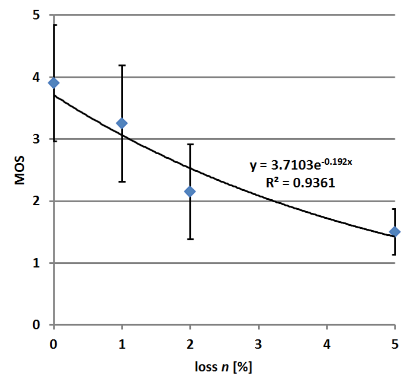
## 6. Conclusions and future work

Given the rising interest in cloud-based thin client services involving access to remote and virtualized PCs, the conducted study has served to provide insight into the actual behavior users engage in when using thin client services, with actual traces collected in an academic setting. Based on statistical traffic analysis and the proposed decision algorithm, passive monitoring of network traffic can be used to determine user behavior and make estimations with regards to network resource requirements. Future work will focus on improving our classification accuracy, and further analyzing additional real world scenarios, such as those focused on use of thin client services in an enterprise/business setting.

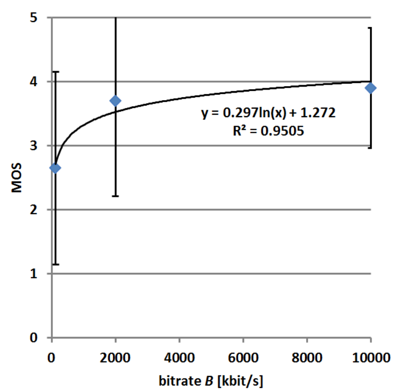
A further focus of this work has been on combining the potential to identify user behavior when using RDC, with the ability to make effective QoE estimations for different behavior categories under various network conditions. Our work has built on previous




(a) MOS vs. delay




(b) MOS vs. loss

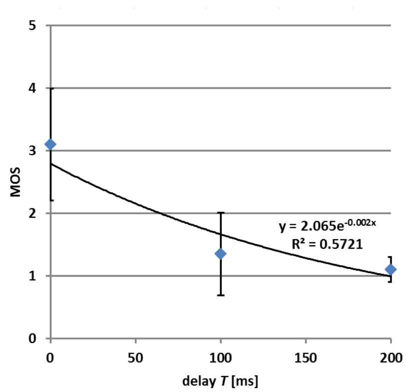



(c) MOS vs. available bandwidth

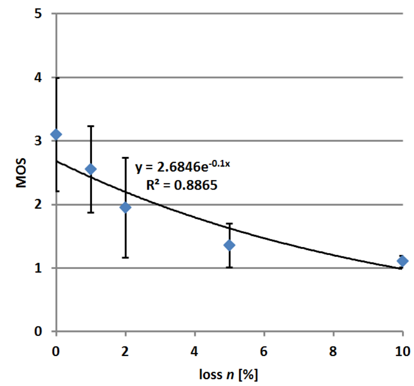
**Fig. 8.** MOS scores for streaming audio shown in relation to (a) delay, (b) loss, and (c) available bandwidth

studies that have explored QoE for RDC services [21, 4, 6], and addressed how behavior detection can be used together with functions modeling QoE in terms of network delay, bandwidth, and loss for different types of remote desktop services. The motivation for such studies lies in the fact that such QoE estimations may then be further used for QoE-driven management mechanisms such as resource planning and dynamic resource (re)allocation in a cloud environment, or service adaptation in order to assure a satisfied customer base. An example of recent work focused on exploiting RDC application identification using statistical mechanisms for the purpose of QoE-driven scheduling is discussed in [1].

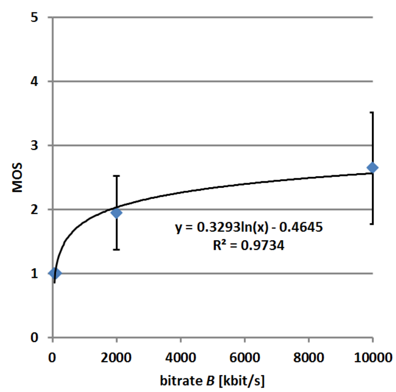
We note that while our studies focused on RDC traffic transmitted using TCP (currently still the most common case), future work in this area should focus also on conducting QoE studies with newer RDC software supporting also UDP (e.g., newer versions of Remote Desktop Protocol available on Windows 8, or VMWare's PCoIP - PC over IP




(a) MOS vs. delay




(b) MOS vs. loss




(c) MOS vs. available bandwidth

**Fig. 9.** MOS scores for streaming video shown in relation to (a) delay, (b) loss, and (c) available bandwidth

protocol) and adaptive graphics, particularly in the context of media streaming. Ongoing developments are focusing on optimizing and adapting the server encoding process in order to improve user perceived quality in light of different network conditions and media requirements.

While we have focused our QoE studies on subjective metrics, objective QoE assessment in terms of e.g., task performance may also be addressed (such as reported in [19, 21, 26, 3]). Furthermore, with the advent of thin client RDC services running on mobile devices using wireless/mobile broadband access, further studies are needed to address QoE in the context of various mobile scenarios. With the expectation that virtual desktop users will increasingly access data-intensive content and multimedia services, such studies will need to address a broader scope of application types than those addressed in this paper.

**Acknowledgments.** This work was partially supported by the Ministry of Science, Education and Sports of the Republic of Croatia and the Ministry of Higher Education and Technology of the Republic of Slovenia under the joint Croatia-Slovenia Cooperation in Science and Technology Project "Services and Applications in the Future Internet: Ensuring Quality and Energy Efficiency". I. Humar's research was also partly performed in the scope of research programme P2-0246 Algorithms and optimization procedures in telecommunications, financed by the Slovenian Research Agency.

## References

1. Arumaithurai, M., Seedorf, J., Dusi, M., Monticelli, E., Lo Cigno, R.: Quality-of-Experience driven Acceleration of Thin Client Connections. In: 12th IEEE International Symposium on Network Computing and Applications (NCA), 2013. pp. 203–210. IEEE (2013)
2. Calyam, P., Patali, R., Berryman, A., Lai, A.M., Ramnath, R.: Utility-directed resource allocation in virtual desktop clouds. Computer networks 55(18), 4112–4130 (2011)
3. Casas, P., Schatz, R.: Quality of experience in cloud services: Survey and measurements. Computer Networks 68, 149–165 (August 2014)
4. Casas, P., Seufert, M., Egger, S., Schatz, R.: Quality of Experience in Remote Virtual Desktop Services. In: 2013 IFIP/IEEE International Symposium on Integrated Network Management (IM 2013). pp. 1352–1357. IEEE (2013)
5. Deboosere, L., Vankeirsbilck, B., Simoens, P., De Turck, F., Dhoedt, B., Demeester, P.: Cloud-based desktop services for thin clients. Internet Computing, IEEE 16(6), 60–67 (2012)
6. Dusi, M., Napolitano, S., Niccolini, S., Longo, S.: A Closer Look at Thin-Client Connections: Statistical Application Identification for QoE Detection. IEEE Communications Magazine 50(11), 195–202 (2012)
7. Emmert, B., Binzenhöfer, A., Schlosser, D., Weiß, M.: Source Traffic Characterization for Thin Client Based Office Applications. In: Dependable and Adaptable Networks and Services, pp. 86–94. Springer (2007)
8. Fiedler, M., Hoßfeld, T., Tran-Gia, P.: A Generic Quantitative Relationship Between Quality of Experience and Quality of Service. Network, IEEE 24(2), 36 –41 (march-april 2010)
9. Hoßfeld, T., Schatz, R., Varela, M., Timmerer, C.: Challenges of qoe management for cloud applications. Communications Magazine, IEEE 50(4), 28–36 (2012)
10. Humar, I., Bester, J., Tomazic, S.: Characterizing Graphical Desktop Sharing System's Workload in Collaborative Virtual Environments. In: Consumer Communications and Networking Conference, 2009. pp. 1–5. IEEE (2009)
11. Humar, I., Pustisek, M., Bester, J.: Evaluating Self-Similar Processes for Modeling Graphical Remote Desktop Systems' Network Traffic. In: 10th International Conference on Telecommunications, 2009. pp. 243–248. IEEE (2009)
12. ITU-T: Rec. P.800.1 - Mean Opinion Score (MOS) terminology (2006)
13. Ivesic, K., Skorin-Kapov, L., Matijasevic, M.: Cross-layer QoE-driven admission control and resource allocation for adaptive multimedia services in LTE. Journal of Network and Computer Applications (2014)
14. Lai, A.M., Nieh, J.: On the Performance of Wide-Area Thin-Client Computing. ACM Transactions on Computer Systems (TOCS) 24(2), 175–209 (2006)
15. Nguyen, T.T., Armitage, G.: A Survey of Techniques for Internet Traffic Classification Using Machine Learning. Communications Surveys & Tutorials, IEEE 10(4), 56–76 (2008)
16. Park, B., Won, Y.J., Choi, M.J., Kim, M.S., Hong, J.W.: Empirical Analysis of Application-Level Traffic Classification Using Supervised Machine Learning. In: Challenges for Next Generation Network Operations and Service Management, pp. 474–477. Springer (2008)
17. Reichl, P., Egger, S., Schatz, R., D'Alconzo, A.: The Logarithmic Nature of QoE and the Role of the Weber-Fechner Law in QoE Assessment. In: ICC. pp. 1–5 (2010)

18. Rugelj, M., Sedlar, U., Volk, M., Sterle, J., Hajdinjak, M., Kos, A.: Novel Cross-Layer QoE-Aware Radio Resource Allocation Algorithms in Multiuser OFDMA Systems. IEEE transactions on communications 62(9), 3196–3208 (2014)
19. Schlosser, D., Staehle, B., Binzenhofer, A., Boder, B.: Improving the qoe of citrix thin client users. In: Communications (ICC), 2010 IEEE International Conference on. pp. 1–6. IEEE (2010)
20. Sen, S., Spatscheck, O., Wang, D.: Accurate, Scalable In-Network Identification of P2P Traffic Using Application Signatures. In: Proceedings of the 13th international conference on World Wide Web. pp. 512–521. ACM (2004)
21. Staehle, B., Binzenhöfer, A., Schlosser, D., Boder, B.: Quantifying the Influence of Network Conditions on the Service Quality Experienced by a Thin Client User. In: Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB), 2008 14th GI/ITG Conference-. pp. 1–15. VDE (2008)
22. Suznjevic, M., Skorin-Kapov, L., Humar, I.: User behavior detection based on statistical traffic analysis for thin client services. In: New Perspectives in Information Systems and Technologies, Volume 2, pp. 247–256. Springer (2014)
23. Tolia, N., Andersen, D.G., Satyanarayanan, M.: Quantifying Interactive User Experience on Thin Clients. Computer 39(3), 46–52 (2006)
24. University of Waikato: WEKA - Waikato Environment for Knowledge Analysis. available at http://www.cs.waikato.ac.nz/ml/weka/
25. Volk, M., Sterle, J., Sedlar, U., Kos, A.: An approach to modeling and control of QoE in next generation networks [Next Generation Telco IT Architectures]. Communications Magazine, IEEE 48(8), 126–135 (2010)
26. Xu, Y., Calyam, P., Welling, D., Mohan, S., Berryman, A., Ramnath, R.: Human-centric composite-quality modeling and assessment for virtual desktop clouds. ZTE Communications 11(1), 27–36 (2013)

**Mirko Suznjevic** received his Ph.D. and B.S. degree in electrical engineering from the University of Zagreb, Faculty of Electrical Engineering and Computing, in 2012 and 2006, respectively. He is currently employed on a post-doc position at the same institution. His research interests are focused on interactive multimedia services, in particular related to massive multiplayer games, Quality of Experience modeling, user behavior monitoring, and statistical network traffic analysis.

**Lea Skorin-Kapov** is an Assistant Professor since 2010 at the University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia. She received her M.Sc. and Ph.D. degrees in electrical engineering from the same University in 2004 and 2007, respectively. From 2001-2009 she was employed as a senior research engineer and project manager in the R&D Center of Ericsson Nikola Tesla d.d., Zagreb, working on projects in collaboration between industry and academia focusing on multimedia QoS negotiation and adaptation. Currently her main research interests are related to QoE modeling of interactive multimedia and Web-based applications, and management/optimization of QoS/QoE. She is a member of IEEE and IEEE standards association, is currently serving as Chapter vice-chair of the IEEE ComSoc Croatia Chapter, and is co-chair for the MMTC QoE Interest Group.

**Iztok Humar** received PhD degrees in the field of Telecommunications at the Faculty of Electrical Engineering (FE) and in the field of Information Management at the Faculty of Economics, University of Ljubljana, Slovenia, in 2007 and 2009 respectively.

He is an Associate Professor at the FE, where he lecturers on design, management and modeling of telecommunication networks. His main research topics include the design, planning and management of telecommunications networks and services, as well as measurement and modeling of network loads and traffic. In 2010 he was Visiting Professor at Huazhong University of Science and Technology (HUST), China, where he worked on energy-efficient wireless networks. Dr. Humar is a senior member of IEEE Communication Society. He served as IEEE Communication Society of Slovenia Chapter Chair, MMTC member and IEEE Slovenia Section secretary.

*Received: October 8, 2014; Accepted: May 15, 2015.*