

Article

Network Traffic Type-Based Quality of Experience (QoE) Assessment for Universal Services

Zheng Xu ¹ and Anguo Zhang ^{2,3,*} 

¹ School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China; bennyxu@126.com

² College of Physics and Information Engineering, Fuzhou University, Fuzhou 350108, China

³ Research Institute of Ruijie, Ruijie Networks Co., Ltd., Fuzhou 350002, China

* Correspondence: anrial@live.cn

Received: 2 September 2019; Accepted: 28 September 2019; Published: 1 October 2019



Abstract: Nowadays, more and more online content providers are offering multiple types of data services. To provide users with a better service experience, Quality of Experience (QoE) has been widely used in the delivery quality measurement of network services. How to accurately measure the QoE score for all types of network services has become a meaningful but difficult problem. To solve this problem, we proposed a unified QoE scoring framework that measures the user experience of almost all types of network services. The framework first uses a machine learning model (random forest) to classify network services, then selects different nonlinear expressions based on the type of service and comprehensively calculates the QoE score through the Quality of Service (QoS) metrics including transmission delay, packet loss rate, and throughput rate. Experiment results show that the proposed method has the ability to be applied on almost all the types of network traffic, and it achieves better QoE assessment accuracy than other works.

Keywords: Quality of Experience; QoE assessment; network traffic classification

1. Introduction

With the growing development of communication and multimedia technology, more and more people use the Internet to obtain information online, such as browsing the web, listening to music, watching videos, downloading files, and so on. In this case, understanding and evaluating the network performance becomes an important task. Traditionally, network operation and maintenance centers use network technology-centric metrics in terms of quality of service (QoS), such as transmission delay, packet loss rate, etc., to measure the network quality. However, it has been demonstrated that the QoS primarily targets at improving the service quality with respect to application-level technical parameters, which lacks sufficient consideration of the user's actual perceptions and feelings [1]. To solve this problem, the user-centric measurement of service quality with the notion of quality of experience (QoE) is introduced, which has drawn much attention in both academia and industry [2–4]. Generally, the assessment of QoE can be done in both subjective and objective ways [5], where subjective evaluation is usually implemented by questionnaires and rating scales [6], can be treated as the more direct and reliable way to evaluate QoE scores. However, it is time-consuming, costly, and inconvenient [7]. Another assessment approach of QoE, objective assessment, uses predefined models to approximate the subjective QoE estimation without human involvement [5].

The existing objective QoE assessment models can be roughly divided into two categories, namely, signal fidelity based and QoS-based. Generally, playback operations are needed when applying the signal fidelity methods, but it is difficult in most streaming services. Thus, the signal fidelity methods cannot be widely applied even if they are highly accurate. In terms of the QoS-based model, the aim is

to find a mapping relation between QoS and QoE. As QoS is described by the network parameters, such as network delay, packet loss rate, and network flow throughput rate; once the mapping relation is built, the QoS-based QoE quantization will be established. Owing to this idea, substantial efforts have been made by engineers and scholars when exploring the QoS-based model, where some works have been done for specific types of network service, such as video stream service [8–12], Internet of Things [13], Instant Messaging [14–16], File Sharing [17], and so on, where proprietary features or parameters are also used in. In [8], the authors proposed a QoE formulation for encrypted YouTube video stream, video source quality-oriented factors like video codec, codec profile, video resolution, bitrate, and video viewing-oriented factors like initial buffering latency, stalling ratio score are both considered into establishing the QoE equation. Chervenets et al. also proposed a QoE evaluation model by using relevant QoS parameters for video service [10], where the QoS formula takes into account the parameters of packet loss, burst level, packet jitter, packet delay, and bandwidth, as well as the weights for each parameters. In [11], a joint framework of QoS and QoE for video transmission over wireless multimedia sensor networks (WMSNs) is proposed, which can support QoS in WMSNs with a lightweight error concealment and utilize the maximum network throughput. The authors of [12] analyzed the streaming video service performance with both the subjective and objective QoS factors (network delay, jitter, and packet loss rate) that affect QoE, which is an inspiration in the processes of both forming QoE model and doing experiments. Li et al. developed a regression equation between QoE and the principal QoS parameters of IoT applications [13].

It should be pointed out that, in the QoE model based on the features of a particular service, it cannot be extended to other network services; however, network operation and maintenance centers often need to ensure that the QoE of all types of network services can be effectively evaluated, thus they can provide users with a better network experience. Therefore, the QoE evaluation model for a single specific service type does not meet the needs of the network operation and maintenance center. To solve this problem, some works have been done on the QoE evaluation algorithms that applicable to multiple service types [18–24]. In [20], Hsu et al. proposed an implementation of QoS/QoE mapping in software-defined networks for ISP to monitor the user-perceived QoE after accessing the service. The authors of [24] also presented some parameteric QoE evaluation for VoIP, online video, FTP services, etc., where only specific part of network parameters are used for different services, such as transmission rate factor in VoIP and rate of correctly received data in FTP. In this paper, we proposed a network traffic classification-based nonlinear mathematical QoE quantization model for universal traffic types, which can be seen as a QoS-based QoE quantization method. For a network data flow, we first classify its traffic type by using a machine learning classifier, then extract its QoS metric values into the corresponding QoE quantization model, after that the QoE score will be obtained. The proposed QoE model has a nonlinear mapping relation with QoS, which will be less influenced by the “marginal effect” than other existing linear mapping quantization approaches. In addition, the proposed model has earned good quantization accuracy and engineering feasibility.

The main contributions of this paper can be listed as follows. (1) An end-to-end QoE score measurement method is proposed. (2) A machine learning classifier (random forest) is used to identify the type of network traffic, where the network traffics are categorized into 11 main types, including on-demand streaming media and live streaming media. (3) More elaborate QoE mathematical quantization formulae of different types of network traffic are given, the way in which different QoS metrics affect QoE is fully considered into the QoE formulas.

The rest of this paper is organized as follows. Section 2 shows some preliminaries about QoS and QoE. In Section 3, we introduce the whole structure of our QoE quantization system, and then the three key procedures, namely, flow features and QoS metrics extraction, random forest classifier, and QoE quantization model, are presented. The experiment details including traffic classification results and QoE quantization results will be given in Section 4. At last, we will summarize the whole paper in Section 5 and give the conclusions.

2. Preliminaries

2.1. Quality of Service (QoS)

QoS was firstly defined by ITU (International Telecommunication Union) in 1994 as a comprehensive measurement of a service’s performance in mobile telecommunication, computer networks, or other digital signal transmission systems. QoS mechanism provides different priorities to different applications, network sessions, or users to guarantee a certain level of performance to a data flow. Most commercial multimedia service providers have applied QoS-oriented optimizing method on effective data flow delivering through resource (bandwidth, throughput, etc.)-limited network, especially for real-time multimedia applications such as IPTV [25], VoIP, online meeting, and multiplayer online games.

QoS mechanism generally uses numeric scores to quantify the network transmission quality of services, which is referred as QoS score. QoS score is usually valued by some related aspects of the network service, like transmission delay, packet loss rate, throughput, network jitter, resource availability, etc. [26]. Typically, a general assessment model is designed as

$$QoS = \sum_{k=1}^N \omega_k \cdot f_k(x_k) \tag{1}$$

where x_k is a network influencing factor, N denotes the number of network influencing factors, ω_k is their corresponding weight, and $f_k(\cdot)$ represents a mapping function.

In the literature, the mapping function is a linear function; however, the linear mapping method ignores the “marginal effect” of some network parameters, which, in fact, cannot fully describe the relation between the network parameters and QoS score; for example, an extreme case, assuming that a certain kind of service is less sensitive to the network transmission delay, namely, it has a relatively small weight in influencing QoS score, and we further assume the network packet loss rate is 100%, which means packet loss will happen with a probability of 100%, with the throughput rate of 100 Mbps. It is clear that the aforementioned network will fail to deliver any packet, and theoretically the QoS scores will rise with the timeline, but unfortunately, a linear mapping function cannot indicate such trend.

2.2. Quality of Experience (QoE)

QoE refers to the user’s assessment of the network service experience under the overall network environment, which is a vital way for content providers and operators to evaluate and improve their service. Generally, we use subjective evaluation–MOS (Mean Opinion Score) to quantify QoE; MOS denotes the levels of experience and can be represented by a single number for evaluation in the range of 1 to 5, as in Table 1.

Table 1. Mean opinion score (MOS) scoring.

Experience	Excellent	Good	Fair	Poor	Bad
Score	5	4	3	2	1

However, obtaining QoE by subjective evaluation method is complex, as well as time-consuming and costly; therefore, realization of real-time quality monitoring is difficult. In order to overcome these difficulties, a series of objective evaluation methods are proposed, via some objective key performance indicators (KPIs) of network service. Remember that the QoE itself is a subjective perceptual score, so one can barely replace it with objective system attributes, directly. What we can do is to describe QoE to a certain degree, with a suitable mapping between objective parameters and QoE.

2.3. Random Forest Classifier

In general, random forest can be seen as a set of unpruned inference trees, who together will vote for the final decision in classification tasks, namely, “majority vote”. Assume a training dataset consisting of N samples with M features in full dimension. Each inference tree in random forest (RF) grows with n ($n < N$) bootstrap samples randomly selected from the same dataset, which can be described as random variable selection; note that only n training samples randomly selected from the dataset is used to build a single tree. Due to the two selection randomness of both the training sample and feature subsets and the advantage of the ensemble learning method, random forest obtains the capability of avoiding the problem of overfitting, high noise immunity, and outstanding computational accuracy. Furthermore, RF can also achieve extremely speedy computation on a parallel architecture, which allows each inference tree to run independently on a CPU core. Because of the excellent classification performance, random forest classifier has been widely used in various applications and tasks.

3. Proposed Method

The proposed method of QoE score measurement, presented in Figure 1, can be divided into three main steps, namely, flow features and QoS metrics extraction, flow traffic classification, and QoE quantization. The features are selected to describe the flow behavior and feed into the “Machine Learning Model” for traffic type classification, QoS metrics (such as time delay, loss rate, and throughput rate) are used in the “QoE Quantization Model” to calculate the QoE score for a network flow with the traffic type identified.

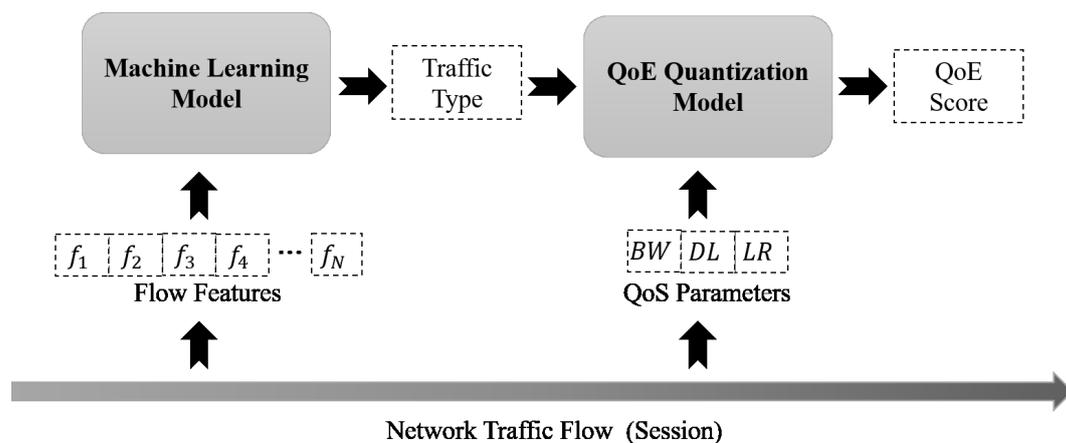


Figure 1. System work flow of our proposed method.

3.1. Flow Features and QoS Metrics Extraction

The network flow features, as shown in Table 2, are extracted as the input of machine learning model, and further labeled to train or validate the model, which can be summarized as the following steps.

- (1) Connect users’ terminal devices (smart phones and tablets) to the wireless network server; here, we use the Wi-Fi hotspot.
- (2) Run the first target APP (application) on the users’ terminal devices, and start to capture the network packet data.
- (3) After a period, e.g., 3 min, restore the captured packet data of the first kind of APP, let users stop running the current APP and open another, collect packet data of the next APP, and repeat such process on each APP one by one until the packet data collection and storage are completed.
- (4) Extract the data features of each kind of APP from its captured data flow.

- (5) Combine the features data of each APP together, as a feature dataset, and divide it into two sub-datasets: training dataset and test dataset.
- (6) Label the training dataset.
- (7) Train the machine learning model with training dataset, and then verify its accuracy with test dataset.

Table 2. Selected features for our proposed random forest-based classification model.

Feature Category	Description	Number of Feature
Packets	bidirectional packet number	2
Bytes	bidirectional Max, Min, Mean, and Var of packet bytes	8
Inter-Arrival Time	bidirectional Max, Min, Mean, and Var of inter-arrival time	8
Total		18

3.2. QoE Quantization Model

Here, we capture ground truth of $\langle QoS \text{ metrics} : QoE \rangle$ data to estimate the coefficients of the QoE quantization model (such as Equations (5) and (7)) of different services, which can be described as a QoE model estimation experiment taking the following steps.

- (1) Connect user terminal devices to the wireless network (the Wi-Fi hotspot).
- (2) Generate ~90–100 different network environments by varying network parameters (DL , LR , and TH) at the Wi-Fi server terminal, where DL , LR , and TH denote the packet delay, loss rate, and flow throughput rate, respectively.
- (3) Collect the QoS data and its corresponding MOS score under different environment, which will then be used as the training/test data and labels.
- (4) Also, here, we need to divide the QoS data into training subset and test dataset without overlap, so the test dataset can help to check and improve model accuracy.
- (5) Label the QoS training dataset and estimate the coefficients through machine learning algorithm.
- (6) Test model accuracy and update the coefficients into a certain extent that meet the practical demand.

To solve the drawback that mentioned in Section 2.2 of the linear mapping function $f_k(\cdot)$, we consider the relations between QoS and network parameters separately.

For the network packet delay, DL ,

$$QoS(DL) = a \cdot \exp(b \cdot DL) + \phi_1 \tag{2}$$

For the packet loss rate, LR ,

$$QoS(LR) = \frac{c}{LR + d} + \phi_2 \tag{3}$$

And for the flow throughput rate, TH ,

$$QoS(TH) = p \cdot \log(TH) + \phi_3 \tag{4}$$

After merging and combining the three influencing factors, we get the QoS expression as follows,

$$\begin{aligned} QoS &= QoS(DL) + QoS(LR) + QoS(TH) \\ &= a \cdot \exp(b \cdot DL) + \frac{c}{LR + d} + p \cdot \log(TH) + q \end{aligned} \tag{5}$$

where DL, LR , and TH denote the transmission delay, packet loss rate, and data throughput rate, respectively. $q = \phi_1 + \phi_2 + \phi_3$ and the parameters a, b, c, d, p, q need to be determined latter for different

types of network services. Markus et al. [27] has demonstrated that when the current QoE value is high, a tiny perturbation in QoS can cause a significant fluctuation, whereas when the current QoE value is low, even a significant change pales; the relation can be described with a formula as follows,

$$\frac{\partial QoE}{\partial QoS} = -(QoE - \gamma) \tag{6}$$

The solution of the above differential formula is

$$QoE = \langle \alpha \cdot \exp(-\beta \cdot QoS) + \gamma \rangle \tag{7}$$

where $\langle \cdot \rangle$ is the rounding function, $\alpha > 0, \beta > 0$ also varies in different types of network traffic, i.e., on-demand video, on-demand audio, live video, live audio, etc. Using Equations (5) and (7), we establish the mathematical formulations between the values of QoE and QoS metrics (DL, LR, TH).

3.3. Determining the Model Parameters

To describe the QoE with mathematical formulas, function fitting methods, such as the least square method and nonuniform rational basis spline (NURBS) fitting method, are usually used to determine the hyperparameter in the formulas. These function fitting methods work well on continuous-value data sets. However, it should be noted that MOS-based QoE scores are discrete integer values in the range of 1 to 5, directly applying these conventional fitting methods on QoE task may cause high computation error. To solve this issue, in this paper, we propose an adaptive hyperparameter calculation algorithm inspired by artificial neural network training method.

Let Ω be the dataset with N samples consists of the network QoS metrics $\langle DL_i, LR_i, TH_i \rangle$ and its labeled QoE score $QoE_i, i = 1, 2, \dots, N$.

According to (7), we define continuous-value variable

$$S = \alpha \cdot \exp(-\beta \cdot QoS) + \gamma \tag{8}$$

where \bar{S}, \underline{S} denotes the upper rounding and down rounding function of S . With $\langle DL_i, LR_i, TH_i \rangle$, we can calculate the corresponding values of S_i, \bar{S}_i and \underline{S}_i .

Then, define the predicted output vector $O_i = [o_{i,1}, o_{i,2}, o_{i,3}, o_{i,4}, o_{i,5}]$, where $o_{\bar{S}_i} = \bar{S}_i - S, o_{\underline{S}_i} = S - \underline{S}$, and the rest of the elements of O are set as 0. Further, define the target output vector $T_i = [t_{i,1}, t_{i,2}, t_{i,3}, t_{i,4}, t_{i,5}]$ where $t_{i,QoE_i} = 1$ and the other elements of T are also set as 0. To set an example, for predicted output $S = 2.4$, and its target output $QoE_i = 3$, we have $\bar{S} = 3, \underline{S} = 2$, then the predicted output vector $O = [0, 0.4, 0.6, 0, 0]$ and the target output vector $T = [0, 0, 1, 0, 0]$.

Define the output error using cross entropy and L2 regularization error:

$$loss_i = - \sum_{k=1}^5 (t_{i,k} \log o_{i,k} + (1 - t_{i,k}) \log (1 - o_{i,k})) + l_2 \cdot \sum_k \Theta_k^2 \tag{9}$$

where l_2 is the regularization factor and Θ denotes the set of all the hyperparameters.

Calculating the gradient of $loss_i$ to model hyperparameters, we have

$$\frac{\partial loss_i}{\partial \alpha} = \frac{\partial loss_i}{\partial O_i} \cdot \frac{\partial O_i}{\partial S_i} \cdot \frac{\partial S_i}{\partial \alpha} = - \sum_{k=1}^5 \left(\left(\frac{t_{i,k}}{o_{i,k}} - \frac{1 - t_{i,k}}{1 - o_{i,k}} \right) \frac{\partial o_{i,k}}{\partial S_i} \cdot \exp(-\beta \cdot QoS_i) \right) \tag{10}$$

Let $g_{i,k} = \left(\frac{t_{i,k}}{o_{i,k}} - \frac{1 - t_{i,k}}{1 - o_{i,k}} \right) \frac{\partial o_{i,k}}{\partial S_i}$, similarly, we can derive

$$\begin{aligned}
 \frac{\partial loss_i}{\partial \beta} &= \frac{\partial loss_i}{\partial O_i} \cdot \frac{\partial O_i}{\partial S_i} \cdot \frac{\partial S_i}{\partial \beta} = \sum_{k=1}^5 (g_{i,k} \cdot \alpha QoS_i \exp(-\beta \cdot QoS_i)) \\
 \frac{\partial loss_i}{\partial \gamma} &= \frac{\partial loss_i}{\partial O_i} \cdot \frac{\partial O_i}{\partial S_i} \cdot \frac{\partial S_i}{\partial \gamma} = - \sum_{k=1}^5 (g_{i,k}) \\
 \frac{\partial loss_i}{\partial a} &= \frac{\partial loss_i}{\partial O_i} \cdot \frac{\partial O_i}{\partial S_i} \cdot \frac{\partial S_i}{\partial QoS_i} \cdot \frac{\partial QoS_i}{\partial a} = \sum_{k=1}^5 (g_{i,k} \cdot \alpha \beta \exp(-\beta \cdot QoS_i + b \cdot DL_i)) \\
 \frac{\partial loss_i}{\partial b} &= \frac{\partial loss_i}{\partial O_i} \cdot \frac{\partial O_i}{\partial S_i} \cdot \frac{\partial S_i}{\partial QoS_i} \cdot \frac{\partial QoS_i}{\partial b} = \sum_{k=1}^5 (g_{i,k} \cdot \alpha \beta a \cdot DL_i \exp(-\beta \cdot QoS_i + b \cdot DL_i)) \\
 \frac{\partial loss_i}{\partial c} &= \frac{\partial loss_i}{\partial O_i} \cdot \frac{\partial O_i}{\partial S_i} \cdot \frac{\partial S_i}{\partial QoS_i} \cdot \frac{\partial QoS_i}{\partial c} = \sum_{k=1}^5 (g_{i,k} \cdot \alpha \beta \exp(-\beta \cdot QoS_i) \cdot \frac{1}{LR_i + d}) \\
 \frac{\partial loss_i}{\partial d} &= \frac{\partial loss_i}{\partial O_i} \cdot \frac{\partial O_i}{\partial S_i} \cdot \frac{\partial S_i}{\partial QoS_i} \cdot \frac{\partial QoS_i}{\partial d} = - \sum_{k=1}^5 (g_{i,k} \cdot \alpha \beta \exp(-\beta \cdot QoS_i) \cdot \frac{c}{(LR_i + d)^2}) \\
 \frac{\partial loss_i}{\partial p} &= \frac{\partial loss_i}{\partial O_i} \cdot \frac{\partial O_i}{\partial S_i} \cdot \frac{\partial S_i}{\partial QoS_i} \cdot \frac{\partial QoS_i}{\partial p} = \sum_{k=1}^5 (g_{i,k} \cdot \alpha \beta \exp(-\beta \cdot QoS_i) \cdot \log(TH_i)) \\
 \frac{\partial loss_i}{\partial q} &= \frac{\partial loss_i}{\partial O_i} \cdot \frac{\partial O_i}{\partial S_i} \cdot \frac{\partial S_i}{\partial QoS_i} \cdot \frac{\partial QoS_i}{\partial q} = \sum_{k=1}^5 (g_{i,k} \cdot \alpha \beta \exp(-\beta \cdot QoS_i))
 \end{aligned} \tag{11}$$

With Equations (10) and (11), we can calculate the gradient of each hyperparameter reflecting to the QoE score fitting error, the hyperparameter determination algorithm is abstracted by the pseudocode presented in Algorithm 1, where parameters bs , eps , and λ are set to 16, 100, and 0.01 in our implementation, respectively.

Algorithm 1: Model Parameter Determination Algorithm.

Input: Dataset $\Omega = \{DL_i, LR_i, TH_i : QoE_i\}, i = 1, \dots, N$.
Output: Model determined hyperparameters: $\Theta = \{\alpha, \beta, \gamma, a, b, c, d, p, q\}$.
 $\Theta = \text{Train}(\Omega)$: Set batch size bs , epoch size eps , and learning rate λ ;
for $i = 1$ **to** eps **do**
 for $j = 1$ **to** $\text{int}(\frac{N}{bs})$ **do**
 Initialize $loss = 0$; **for each sample** $\{DL_k, LR_k, TH_k : QoE_k\}$ **in** $\Omega[bs * (j - 1) + 1 : bs * j]$ **do**
 | According to (9), $loss += loss_k$;
 end
 Calculate average loss by $\overline{loss} = \frac{loss}{bs}$; **for each hyperparameter** θ_k **in** Θ **do**
 | According to (10) and (11), update θ_k by gradient: $\theta_k \leftarrow \theta_k - \lambda \cdot \frac{\partial \overline{loss}}{\partial \theta_k}$
 end
 end
end
return Θ

4. Results

4.1. Data Collection

To verify our proposed method, we select 11 different types of traffic service: on-demand video, on-demand audio, live video, live audio, web browsing, instant message, file sharing, online game, location, meeting video, and voice call. Here, experiments are conducted with the ground truth

data captured from internet with the “.pcap” file format. The data is collected from a university campus, for each type of network traffic, 9-11 different corresponding APPs are used. Details of the collected network flow data has been presented in Table 3, where “APP Number” denotes the number of different APPs used for a service type and “File Number” denotes the number of the collected “.pcap” files.

Table 3. Details of APP flow data collection.

	APP Number	File Number	Total Size
on-demand video	10	357	17.9 GB
on-demand audio	10	285	4.05 GB
live video	9	300	16.6 GB
live audio	11	354	1.41 GB
web browsing	10	368	2.87 GB
instant message	10	342	0.41 GB
file sharing	10	331	12.49 GB
online game	10	194	0.95 GB
location	10	351	0.56 GB
meeting video	10	244	9.12 GB
voice call	10	255	2.47 GB
Total	110	3781	67.83 GB

4.2. Traffic Classification Performance

The confusion matrix of classification results is presented in Figure 2, where the tick label annotations are OA: on-demand audio; WB: web browsing; IM: instant message; FS: file sharing; OG: online game; LV: live video; LO: location; LA: live audio; OV: on-demand video; MV: meeting video; and VC: voice call. Further, Table 4 shows the Precision, Recall Rate, and F1 Score of all the network service types, where Precision, Recall Rate, and F1 Score are defined as

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

$$Recall = \frac{TP}{TP + FN} \quad (13)$$

$$F1Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (14)$$

where TP (True Positive) is the number of instances correctly classified as X, FP (False Positive) is the number of instances incorrectly classified as X, and FN (False Negative) is the number of instances incorrectly classified as Not-X. Precision means the proportion of the true positive samples that are determined by the classifier to be positive; the recall rate reflects the proportion of true positive samples that are correctly determined to the total positive samples. The F1 score is the harmonic average of the precision and the recall rate, when only both metrics are high, the F1 score will achieve a high value.

On the test dataset of almost all network traffic types, our approach can achieve a competitive average classification accuracy of nearly 90%. From Figure 2, we find that the misclassification error rate between voice call and meeting video is significantly higher than that of other services; the most likely reason is that some of the voice call and meeting video use the same network transmission protocols, such as RTP, H.323, and SIP. Moreover, in some meeting video services, the voice call is integrated as one of the basic services, which means the two service flows have overlaps, thus higher classification error occurs.

Table 4. Classification performance of all the network traffic types.

	Precision	Recall	F1-Score
on-demand audio	87.98%	79.65%	83.61%
web browsing	87.68%	81.25%	84.34%
instant message	98.34%	95.69%	96.99%
file sharing	91.49%	90.94%	91.21%
online game	87.10%	83.51%	85.26%
live video	91.03%	91.33%	91.18%
location	92.02%	95.56%	93.79%
live audio	85.19%	90.96%	87.98%
on-demand video	75.57%	84.03%	79.58%
meeting video	73.80%	69.26%	71.46%
voice call	66.67%	78.43%	72.07%
Average	88.90%	90.11%	89.51%

We compare the classification performance between our applied random forest and other popular classifiers, e.g., support vector machine (SVM), feedforward neural network (NN), and top-K nearest neighbors (top1-NN and top5-NN); the results shown in Figure 3 suggest that the random forest classifier outperforms the other classifiers in Precision, Recall Rate, and F1-Score.

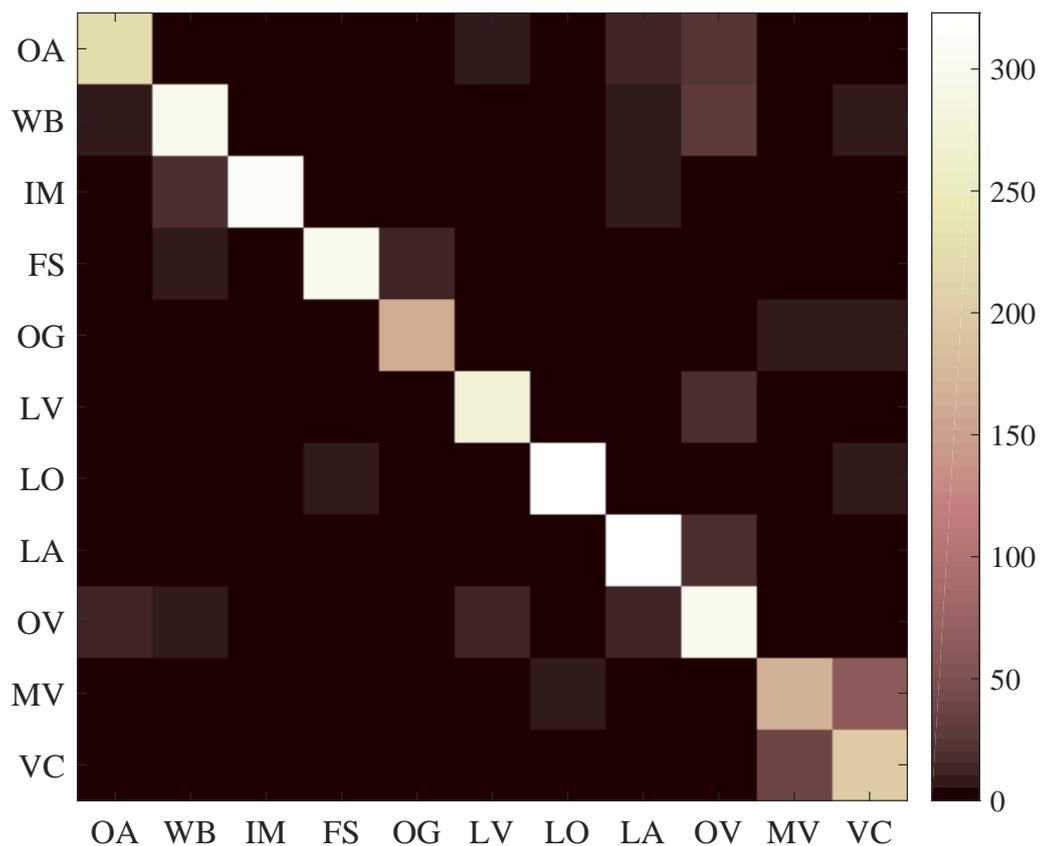


Figure 2. Confusion matrix of the classification results.

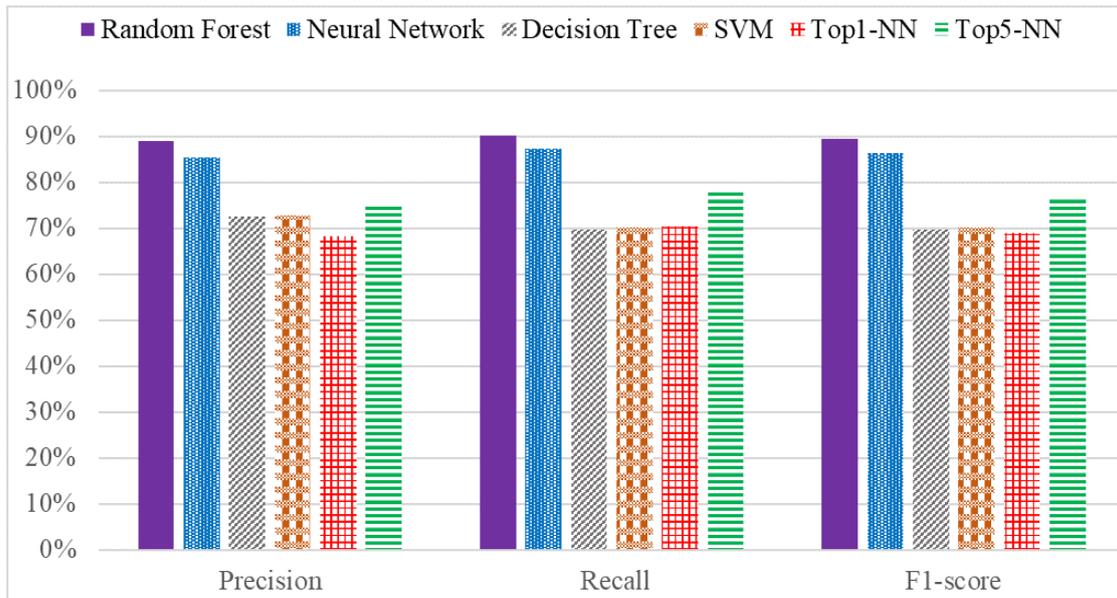


Figure 3. Traffic classification performance comparison between our applied random forest method and other classifiers.

4.3. QoE Quantization Performance

In Tables 5 and 6, the central subfigures present the scores of the labeled and predicted QoE scores of the different network services in the test scenario, and the traffic model hyperparameters are also presented. Meanwhile, the central subfigures also provide an insight into the absolute prediction errors for test scenarios. Additionally, normalized root-mean-squared error (NRMSE) is used to measure error between the model predicted MOS score and the ground truth MOS score, NRMSE is calculated by

$$NRMSE = \frac{1}{N} \sqrt{\sum (QoE_p - QoE_T)^2} \tag{15}$$

where N denotes the number of test samples, QoE_p and QoE_T are the model predicted QoE and the ground truth QoE value, respectively. In Table 7, we compare the NRMSE of our model with the other five works. It can be seen that among the 11 types of traffics, the model we propose achieves the seven smallest values of NRMSE, whereas the difference of the other four NRMSE with minimum values are not large (0.1581 vs. 0.1327 [28] in file sharing, 0.1768 vs. 0.1699 [29] in live video, 0.1803 vs. 0.1722 [24] in live audio, 0.2151 vs. 0.1856 [24] in meeting video).

Further, one may notice that the NRMSEs are lower in on-demand multimedia services than those in live multimedia, online games, and meeting video services. It can be interpreted that, nowadays, on-demand steaming services have a much higher proportion than live streaming services, and people are familiar with them. Therefore, users' will have consistent senses toward on-demand streaming services, which allows a higher quantization accuracy with the mathematical model. In addition, to improve the accuracy of MOS prediction, we should ensure that the QoS metric range is fully covered by the labeled dataset and cross validation methods can be applied to check the MOS inconsistency caused by individual's subjective fluctuations.

Table 5. Flow characteristics of different network service types. For each type of network service, we selected three APP packet streams to demonstrate the flow distribution over time.

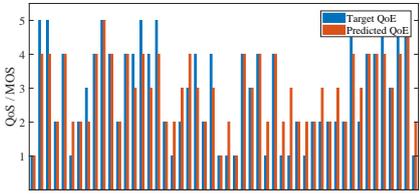
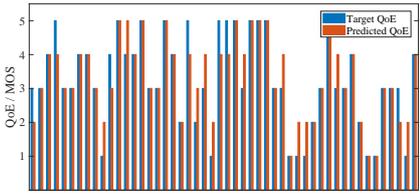
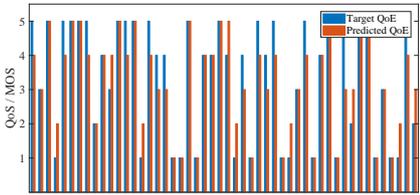
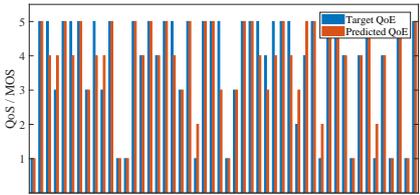
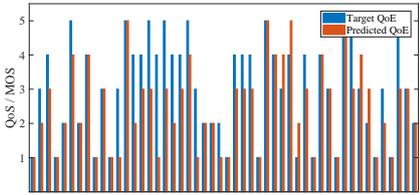
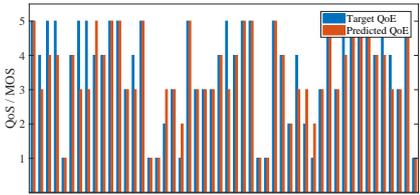
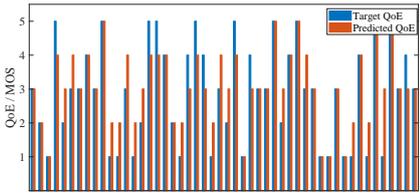
Service Type	Labeled QoE vs. Predicted QoE	Parameter Values
on-demand audio		$\alpha = 1.1466, \beta = 0.0299,$ $\gamma = -0.3209,$ $a = 7.6178, b = 0.5831,$ $c = -15.2617, d = 0.6199,$ $p = -12.1927, q = -0.0103$
web browsing		$\alpha = 0.5003, \beta = 13.0852,$ $\gamma = -0.0488,$ $a = 0.0136, b = 0.9295,$ $c = -0.0420, d = 0.4487,$ $p = -0.0182, q = -0.0703$
instant message		$\alpha = 0.2375, \beta = 1.2047,$ $\gamma = 2.0718,$ $a = 0.3321, b = 1.2422,$ $c = -0.4628, d = 0.3522,$ $p = -0.0937, q = -1.0115$
file sharing		$\alpha = 1.4757, \beta = 0.4038,$ $\gamma = 2.0749,$ $a = 0.5913, b = 1.1134,$ $c = -1.2989, d = 0.4902,$ $p = -0.3592, q = 0.4521$
online game		$\alpha = 0.8771, \beta = 1.9647,$ $\gamma = 0.5635,$ $a = 0.2743, b = 0.9683,$ $c = -0.0212, d = -0.8657,$ $p = -0.0033, q = 0.0893$
live video		$\alpha = 1.6466, \beta = 0.0293,$ $\gamma = 0.1270,$ $a = 4.3215, b = 0.7696,$ $c = -14.1519, d = 0.4361,$ $p = -8.6823, q = 0.0838$
location		$\alpha = 1.0901, \beta = 0.2868,$ $\gamma = 1.5876,$ $a = 0.7781, b = 1.0480,$ $c = -1.9215, d = 0.4574,$ $p = -0.1893, q = 0.5047$

Table 6. Flow characteristics of different network service types. For each type of network service, we selected three APP packet streams to demonstrate the flow distribution over time.

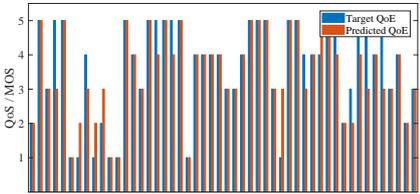
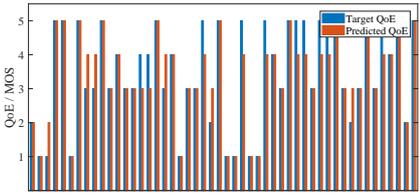
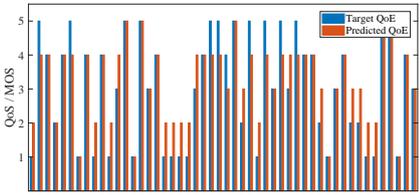
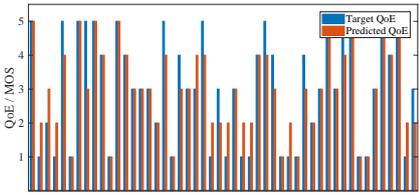
Service Type	Labeled QoE vs. Predicted QoE	Parameter Values
live audio		$\alpha = 1.1085, \beta = 0.7593,$ $\gamma = 1.5062,$ $a = 0.3965, b = 1.0749,$ $c = -0.8575, d = -0.7646,$ $p = -0.0445, q = -2.7661$
on-demand video		$\alpha = 1.3566, \beta = 0.0424,$ $\gamma = -0.0199,$ $a = 2.8019, b = 0.8041,$ $c = -8.3946, d = 0.2898,$ $p = -5.6192, q = 0.3479$
meeting video		$\alpha = 1.1468, \beta = 0.2167,$ $\gamma = 1.8257,$ $a = 0.7221, b = 1.1079,$ $c = -1.4156, d = 0.1628,$ $p = -1.9646, q = 1.8747$
voice call		$\alpha = 1.2947, \beta = 0.1943,$ $\gamma = 0.8373,$ $a = 1.1351, b = 1.2000,$ $c = -2.6529, d = 0.4672,$ $p = -0.4855, q = 1.0115$

Table 7. Comparison of Quality of Experience (QoE) assessment performance with other works, in terms of normalized root-mean-squared error (NRMSE). For each type of traffic, the minimum NRMSE value is marked in bold.

Ref	On-Demand Audio	Web Browsing	Instant Messaging	File Sharing	Online Game	Live Video
[30]	0.3213	0.2122	0.3198	0.2355	0.3476	0.2316
[31]	0.1654	0.1922	0.1909	0.1327	0.2728	0.2061
[29]	0.1662	/	/	/	/	0.1699
[18]	0.1688	/	/	/	/	0.2715
[24]	0.1529	0.1688	/	0.1821	/	0.1933
ours	0.1471	0.1500	0.1803	0.1581	0.2669	0.1768
Ref	Location	Live Audio	On-Demand Video	Meeting Video	Voice Call	
[30]	0.2527	0.2961	0.1420	0.3092	0.1983	
[31]	0.2314	0.1738	0.1785	0.1925	0.2082	
[29]	/	/	0.1538	0.2363	/	
[18]	/	0.1966	0.1689	0.2209	0.1735	
[24]	/	0.1722	0.1382	0.1856	0.1775	
ours	0.1871	0.1803	0.1369	0.2151	0.1696	

4.4. Sensitivity Analysis of QoS Metrics

To verify the rationality of QoE quantization models, we analyze the QoE sensitivities of different services over the fluctuations of QoS metrics (*DL*, *LR*, and *TH*).

4.4.1. Time Delay

The sensitivity of QoE to time delay can be calculated by

$$\Psi_{DL} = \frac{\partial QoE}{\partial DL} = -\alpha\beta ab \cdot \exp(-\beta \cdot QoS + b \cdot DL) \tag{16}$$

We assume that DL lays in the range of $[0, 1000]$, meaning the time delay is between 0 and 1000 milliseconds. The sensitivity curves of five services are drawn in Figure 4, and it can be seen that their absolute values have the following relation, $\{|\Psi_{DL}|(IM), |\Psi_{DL}|(VC), |\Psi_{DL}|(LA), |\Psi_{DL}|(OG)\} > \{\Psi_{DL}(LO), |\Psi_{DL}|(MV), |\Psi_{DL}|(LV)\} > \{|\Psi_{DL}|(OV), |\Psi_{DL}|(OA), |\Psi_{DL}|(FS), |\Psi_{DL}|(WB)\}$. Also, note that in the range $[0, 400]$ ms of DL , the on-demand video service and on-demand audio service curves are approximately coincident, which lead to the following conclusions.

- (1) The values of all the curves in the entire time delay analysis domain (0 to 1000 ms) are less than 0, indicating that the increase of time delay has a negative effect on all QoEs of the network services, and will lead to the attenuation of QoE.
- (2) The QoE of IM, VC, LA, and OG is the most susceptible to time delay, followed by LO, MV, and LV. Live services are most sensitive to time delay, whereas other traffic, like OV, OA, FS, and WB services, are the least affected, which correctly matched with our intuitive experience. For instance, web pages opening one second earlier will lead to the same reading experience as the one opened a second later, but people will certainly feel bad for the buffer time while watching live video.
- (3) In the range of 0 to 400 ms, the effect of time delay on the on-demand services (on-demand video and on-demand audio) is almost the same, but when the time delay continues to increase, the QoE deterioration of on-demand video is more pronounced. A reasonable explanation to this phenomenon is the time-out retransmission mechanism. The amount of data that the video service has to retransmit is larger than other services, and due to the limitation of the throughput rate, it is unable to meet a high-load data transmission.

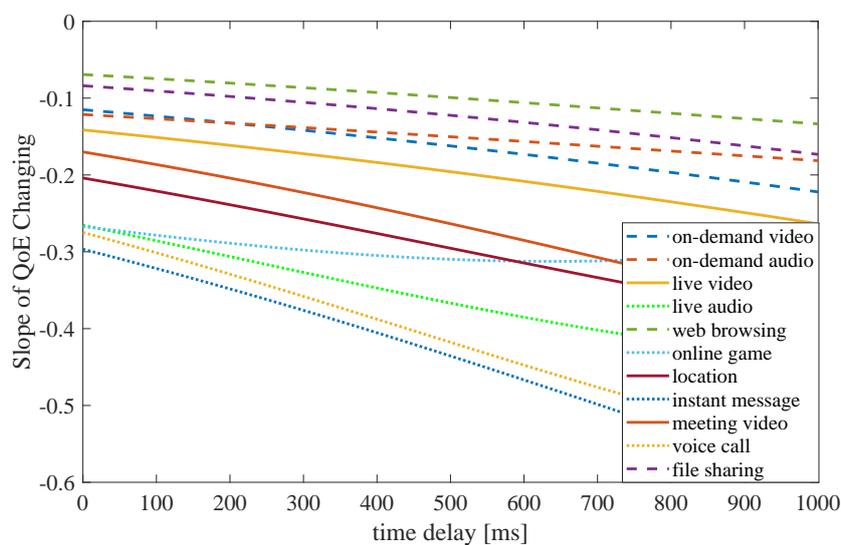


Figure 4. Effects of the packet time delay on QoE attenuating.

4.4.2. Loss Rate

The sensitivity of QoE to packet loss rate (LR) is

$$\Psi_{LR} = \frac{\partial QoE}{\partial LR} = \frac{\alpha\beta c}{(LR + d)^2} \cdot \exp(-\beta \cdot QoS) \tag{17}$$

Here, we assume that LR is between 0 and 50%, the QoE changing slope values in Figure 5 can be described as $\{|\Psi_{LR}|(MV), |\Psi_{LR}|(OV)\} > \{|\Psi_{LR}|(FS), |\Psi_{LR}|(LO), |\Psi_{LR}|(LV), |\Psi_{LR}|(VC)\} > \{|\Psi_{LR}|(WB), |\Psi_{LR}|(OA), |\Psi_{LR}|(LA), |\Psi_{LR}|(OG)\}$. Further, we can draw the following conclusions.

- (1) The values of all the curves are less than 0 in the entire packet loss rate analysis domain, which indicates that the increasing packet loss rate has a negative effect on all QoEs of the data service, which will lead to the attenuation of QoE.
- (2) For most types of data services, in the range of [0, 10%], a slight increase in the packet loss rate leads to a sharp drop in QoE. When the packet loss rate reaches to a relative high level, the impact on QoE will be less obvious than it in the range of 0 – 20%, which is mainly reflected in the range of 20 to 30%.
- (3) From high to low order, it can be seen that the more data to be transmitted, the less dependable a certain kind of service is. As example, live audio requires less data flow due to its high compression rate.

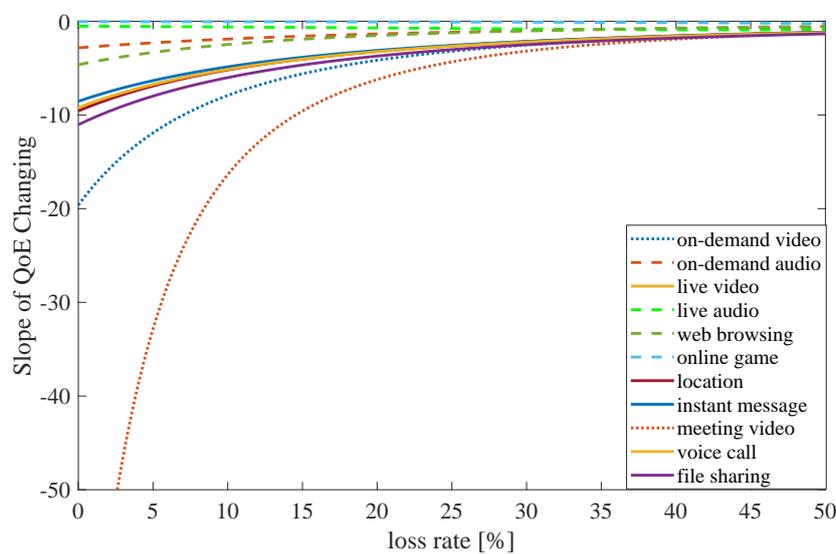


Figure 5. Effects of the packet loss rate on QoE attenuating.

4.4.3. Throughput Rate

QoE sensitivity to throughput rate (TH):

$$\Psi_{TH} = \frac{\partial QoE}{\partial TH} = -\frac{\alpha\beta p}{TH} \cdot \exp(-\beta \cdot QoS) \tag{18}$$

We assume that TH takes the range from 0 to 500 Kbps, the QoE changing slope values in Figure 6 can be described as $\{\Psi_{TH}(LV), \Psi_{TH}(OV)\} > \{\Psi_{TH}(MV), \Psi_{TH}(OA), \Psi_{TH}(WB), \Psi_{TH}(LA), \Psi_{TH}(FS)\} > \{\Psi_{TH}(VC), \Psi_{TH}(IM), \Psi_{TH}(LO), \Psi_{TH}(OG)\}$. Further, we can make the following conclusions:

- (1) The values on the curves are greater than 0 in the entire analysis domain, indicating that the increase in throughput rate has a positive effect on the QoE of all network services.
- (2) At the initial stage of the throughput rate, the values of curves are relatively large, but then drop quickly to ~0, except in network traffic of the video service. This implies that for all network services, QoE will increase significantly when the throughput rate starts to increase from 0. However, the increase in QoE will become less obvious when the throughput rate reaches a certain level. Therefore, we can conclude that increasing throughput rate will not lead to increasing QoE, as long as the throughput rate meeting the service requirement.

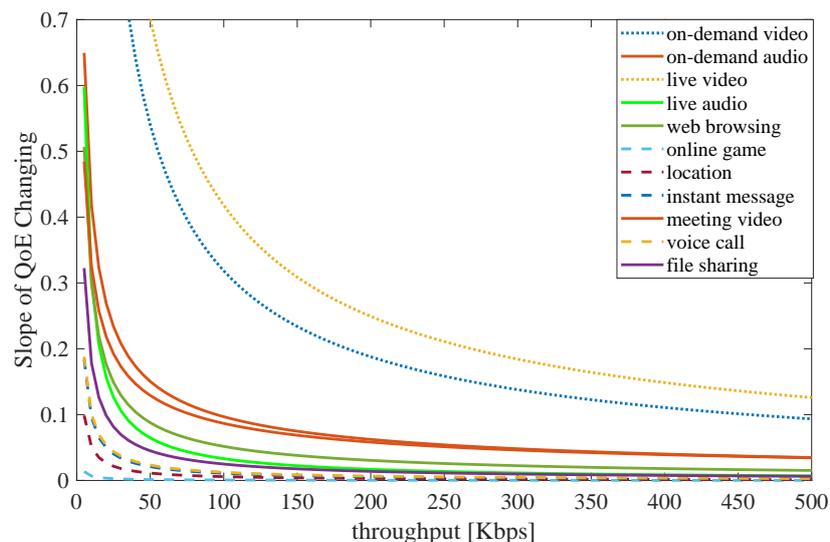


Figure 6. Effects of the flow throughput rate on QoE attenuating.

5. Conclusions

In this paper, we proposed a QoE quantization method based on network traffic classification and nonlinear mathematical model. For network packet flow, we first identify the flow type provided by this session through a random forest model. Simultaneously, network QoS metrics are also measured, including time delay, packet loss rate, and throughput rate. According to the service type of the session flow and its corresponding QoS metrics, we use the QoE description mathematical model to quantify its QoE score. Experiment results show that compared with other works, our proposed method is more suitable for almost all the types of network traffic and further achieves better measurement accuracy. On the other hand, the QoE sensitivity to time delay, packet loss rate and throughput rate are also analyzed, respectively. Results suggest that the QoE sensitivity to QoS metrics varies with the types of network traffics according to the model parameters. Our future research will consider using some network traffic specific features to measure the QoE score in addition to the QoS metrics.

Author Contributions: A.Z. conceived the idea and designed and organized the information in this paper. Z.X. and A.Z. wrote and reviewed the manuscript and contributed to changes in all sections. Z.X. programmed the experiments. All authors read and approved the final manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Staelens, N.; Moens, S.; den Broeck, W.V.; Mariën, I.; Vermeulen, B. Assessing quality of experience of IPTV and video on demand services in real-life environments. *IEEE Trans. Broadcast.* **2010**, *56*, 458–466. [[CrossRef](#)]
2. Chen, L.; Deng, X.; Qi, Y.; Mohammed, S.A.; Fei, Z.; Zhang, J. A Method to Evaluate Quality of Real Time Video Streaming in Radio Access Network. In Proceedings of the IEEE International Conference on Computer and Information Technology, Xi'an, China, 11–13 September 2014; pp. 340–344.
3. Bouten, N.; Latré, S.; de Meerssche, W.V.; Vleeschauwer, B.D.; Schepper, K.D.; Leekwijck, W.V.; Turck, F.D. A Multicast-Enabled Delivery Framework for QoE Assurance of Over-The-Top Services in Multimedia Access Networks. *Netw. Syst. Manag.* **2013**, *21*, 677–706. [[CrossRef](#)]
4. Duanmu, Z.; Zeng, K.; Ma, K.; Rehman, A.; Wang, Z. A Quality-of-Experience Index for Streaming Video. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 154–166. [[CrossRef](#)]
5. Brooks, P.; Hestnes, B. User measures of quality of experience: Why being objective and quantitative is important. *IEEE Netw.* **2010**, *24*, 8–13. [[CrossRef](#)]
6. Jain, R. Quality of experience. *IEEE Multimed.* **2004**, *11*, 95–96. [[CrossRef](#)]

7. Claypool, M.; Le, P.; Wased, M.; Brown, D. Implicit interest indicators. In Proceedings of the ACM International Conference on Intelligent User Interfaces, Santa Fe, NM, USA, 14–17 January 2001; pp. 33–40.
8. Pan, W.; Cheng, G. QoE Assessment of Encrypted YouTube Adaptive Streaming for Energy Saving in Smart Cities. *IEEE Access* **2018**, *6*, 25142–25156. [[CrossRef](#)]
9. Barman, N.; Martini, M. QoE Modeling for HTTP Adaptive Video Streaming: A Survey and Open Challenges. *IEEE Access* **2019**, *7*, 30831–30859. [[CrossRef](#)]
10. Chervenets, V.; Romanchuk, V.; Beshlcy, H.; Khudyy, A. QoS/QoE Correlation Modified Model for QoE Evaluation on Video Service. In Proceedings of the International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science, Lviv, Ukraine, 23–26 February 2016; pp. 664–666.
11. Usman, M.; Yang, N.; Jan, M.A.; He, X.; Xu, M.; Lam, K.M. A Joint Framework for QoS and QoE for Video Transmission over Wireless Multimedia Sensor Networks. *IEEE Trans. Mob. Comput.* **2018**, *17*, 746–759. [[CrossRef](#)]
12. Ning, Z.; Liu, Y.; Wang, X.; Feng, Y.; Kong, X. A Novel QoS-Based QoE Evaluation Method for Streaming Video Service. In Proceedings of the IEEE International Conference on Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data, Exeter, UK, 21–23 June 2017; pp. 956–961.
13. Li, L.; Rong, M.; Zhang, G. An Internet of Things QoE Evaluation Method Based on Multiple Linear Regression Analysis. In Proceedings of the International Conference on Computer Science & Education, Cambridge, UK, 22–24 July 2015.
14. Alsulami, M.M.; Al-Aama, A.Y. Exploring User’s Perception of Storage Management Features in Instant Messaging Applications: A Case on WhatsApp Messengere. In Proceedings of the International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 1–3 May 2019 .
15. Xin, X.; Wang, W.; Huang, A.; Shan, H. Quality of experience modeling and fair scheduling for instant messaging service. In Proceedings of the International Conference on Wireless Communications and Signal Processing, Hefei, China, 23–25 October 2014.
16. Xin, X.; Wang, W.; Huang, A.; Shan, H. Representative service based quality of experience modeling for instant messaging service. In Proceedings of the IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC), Washington, DC, USA, 2–5 September 2014 .
17. Hopfeld, T.; Skorin-Kapov, L.; Heegaard, P.E.; Varela, M. Definition of QoE Fairness in Shared Systems. *IEEE Commun. Lett.* **2017**, *21*, 184–187. [[CrossRef](#)]
18. Hsu, W.; Lo, C. QoS/QoE Mapping and Adjustment Model in the Cloud-based Multimedia Infrastructure. *IEEE Syst. J.* **2014**, *8*, 247–255. [[CrossRef](#)]
19. Mitra, K.; Zaslavsky, A.; Ahlund, C. Context-Aware QoE Modelling, Measurement, and Prediction in Mobile Computing Systems. *IEEE Trans. Mob. Comput.* **2015**, *14*, 920–936. [[CrossRef](#)]
20. Hsu, W.; Wang, X.; Yeh, S.; Huang, P. The Implementation of a QoS/QoE Mapping and Adjusting Application in Software-Defined Networks. In Proceedings of the International Conference on Intelligent Green Building and Smart Grid, Prague, Czech Republic, 27–29 June 2016.
21. Aguilera, N.; Bustos, J.; Lalane, F. Adkintun Mobile: Study of Relation between QoS and QoE in Mobile Networks. *IEEE Lat. Am. Trans.* **2016**, *14*, 2770–2772. [[CrossRef](#)]
22. Chang, H.; Hsu, C.F.; Hopfeld, T.; Chen, K. Active Learning for Crowdsourced QoE Modeling. *IEEE Trans. Multimed.* **2018**, *20*, 3337–3352. [[CrossRef](#)]
23. Zhou, L.; Wu, D.; Wei, X.; Dong, Z. Seeing Isn’t Believing: QoE Evaluation for Privacy-Aware Users. *IEEE J. Sel. Areas Commun.* **2019**, *37*, 1656–1665. [[CrossRef](#)]
24. Tsolkas, D.; Liotou, E.; Passas, N.; Merakos, L. A survey on parametric QoE estimation for popular services. *J. Netw. Comput. Appl.* **2017**, *77*, 1–17. [[CrossRef](#)]
25. Olariu, C.; Song, Y.; Brennan, R.; McDonagh, P.; Hava, A.; Thorpe, C.; Murphy, J.; Murphy, L.; French, P.; de Fréin, R. Integration of QoS Metrics, Rules and Semantic Uplift for Advanced IPTV Monitoring. *J. Netw. Syst. Manag.* **2015**, *23*, 673–708.
26. Davy, A.; Jennings, B.; Botvich, D. QoSPlan: A Measurement Based Quality of Service aware Network Planning Framework. *J. Netw. Syst. Manag.* **2013**, *21*, 474–509. [[CrossRef](#)]
27. Markus, F.; Tobias, H.; Phuoc, T. A generic quantitative relationship between Quality of experience and Quality of Service. *IEEE Netw.* **2010**, *24*, 36–41.

28. 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 Proceedings of the IEEE International Conference on Communications, Cape Town, South Africa, 23–27 May 2010.
29. Kim, H.; Choi, S. QoE assessment model for multimedia streaming services using QoS parameters. *Multimed. Tools Appl.* **2014**, *72*, 2163–2175. [[CrossRef](#)]
30. Kim, H.; Lee, D.; Lee, J.; Lee, K.; Lyu, W.; Choi, S. The QoE Evaluation Method through the QoS-QoE Correlation Model. In Proceedings of the Fourth International Conference on Networked Computing and Advanced Information Management, Gyeongju, Korea, 2–4 September 2008.
31. Reichl, P.; Tuffin, B.; Schatz, R. Logarithmic Laws in Service Quality Perception: Where Microeconomics Meets Psychophysics and Quality of Experience. *Telecommun. Syst.* **2013**, *52*, 1–14. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).