CrossMark

Data analysis on video streaming QoE over mobile networks

Qingyong Wang[1], Hong-Ning Dai[1*] , Di Wu[2,3] and Hong Xiao[4]

Abstract

One of recent proposals on standardizing quality of user experience (QoE) of video streaming over mobile network is video Mean Opinion Score (vMOS), which can model QoE of video streaming in 5 discrete grades. However, there are few studies on quantifying vMOS and investigating the relationship between vMOS and other quality of service (QoS) parameters. In this paper, we address this concern by proposing a novel data analytical framework based on video streaming QoE data. In particular, our analytical model consists of *K*-means clustering and logistic regression. This model integrates the benefits of both these two models. Moreover, we conduct extensive experiments on realistic dataset and verify the accuracy of our proposed model. The results show that our proposed framework outperforms other existing methods in terms of prediction accuracy. Moreover, our results also show that vMOS is essentially affected by many QoS parameters such as initial buffering latency, stalling ratio, and stalling times. Our results offer a number of insights in improving QoE of video streaming over mobile networks.

Keywords: Quality of service (QoS), Quality of experience (QoE), *K*-means, Logistics regression, Video streaming, Mobile networks

1 Introduction

Video streaming is becoming one of the most popular services over mobile networks. It is predicted in [1] that the traffic caused by video streaming will occupy more than 77% of all consumer Internet traffic by 2021, among which mobile video traffic will be more than 55% of all video traffic. The growing demands on video streaming over mobile networks inevitably lead to the challenges in optimizing network resource in order to improve the user *perceptual* experience. Many previous studies mainly focus on improving quality of service (QoS) of video streaming over mobile networks. Typical QoS measures include throughput, bandwidth, outage, jitter, and delay [2]. However, most of these QoS metrics fail to characterize user perceptual experience, which is also called quality of experience (QoE). It is more crucial to conduct video quality assessment from QoE than that from QoS [3, 4] because (i) enhancing QoS does not directly improve QoE [5] and (ii) only improving QoS sometimes significantly increases operating expenditure, consequently decreasing the profit of service providers [6].

Therefore, QoE improvement of video streaming over mobile networks has received extensive attention recently. In particular, the work of [7] investigates QoE-driven cross-layer optimization for video transmission in wireless networks. Ramamurthi et al. [8] propose a resource management scheme at network core in wireless networks to improve video QoE. The work of [9] presents a large-scale measurement-based study on the effects of Internet path selection in video QoE and offers several QoE enhancement schemes.

However, the prerequisite of QoE improvement of video streaming is to quantify QoE appropriately. Video QoE assessment schemes can be generally categorized into *subjective tests*, *objective assessments*, and *data-driven analysis* [3]. Compared with subjective tests and objective assessments, data-driven analysis is more promising due to the availability of massive datasets and the accuracy of characterizing user perception while overcoming the drawbacks of subjective tests and objective assessments (such as high cost and insufficient human visual system knowledge). In particular, the work of [6] proposes a data-driven model to quantify the metrics affecting video

*Correspondence: hndai@ieee.org
[1]Faculty of Information Technology, Macau University of Science and Technology, Wai Long Avenida, Macau SAR, China
Full list of author information is available at the end of the article

Wang *et al. EURASIP Journal on Wireless Communications and Networking* (2018) 2018:173

Page 2 of 10

QoE. Jiang et al. [10] improve video QoE by exploiting data-driven QoE prediction. The work of [11] improves the video bitrate adaptation based on data-driven QoE prediction. Huang et al. [12] propose a dynamic adaptive streaming via HTTP to optimize user QoE. Ref. [13] presents a Multi-Constraint Quality-of-Experience (QoE) centric Routing (MCQR) scheme to improve user video streaming QoE over mobile networks. The work of [14] presents a queue-based model to analyze the video buffer using discrete-time analysis. Marai et al. [15] propose a client-server cooperation-based approach to achieve efficiency, fairness, and stability of video adaptive streaming.

In addition to the above efforts, there are also other solutions on standardizing QoE. One of recent video QoE measurement standards is U-vMOS (User/Unified/Ubiquitous video Mean Opinion Score)[1], which was proposed by Huawei in 2016 [16]. The score of vMOS is essentially established according to Mean Opinion Score (MOS) standardized by International Telecommunication Union (ITU) [17] as shown in Table 1, where discrete grades from 1 to 5 represent bad, poor, fair, good, and excellent, respectively. It is shown in [16] that vMOS at video playback startup is mainly determined by three key factors: *video quality*, *initial buffering delay*, and *video freezing duration*, each of which is also affected by multiple QoS variables. Recently, Pan et al. [18] investigate machine learning-based bitrate estimation on YouTube video streaming based on Huawei's vMOS assessment model. However, they just give a mathematical expression of vMOS based on their subjective estimations. The work of [19] presents using vMOS to investigate the impacts of several system parameters on long-term evolution (LTE) networks based on simulations. *To the best of our knowledge, there is no data-driven QoE analysis on vMOS.*

Therefore, this paper aims to conduct data-driven QoE analysis on vMOS. In particular, we obtain a realistic dataset on video QoE based on SpeedVideo Global Operating Platform (SVGOP) established by Huawei. This dataset has the following unique characteristics: (1) heterogeneous data types, (2) positive/negative correlations, and (3) dependence of features; these characteristics result in the difficulties in analyzing video QoE data.

To address the above concerns, we propose a data-driven analysis framework to analyze the relationship between vMOS and other QoS parameters. Although our previous work [20] presented preliminary results on quantifying vMOS and other QoS parameters, this study is significantly different from our previous work in the following aspects: (1) we conduct a data preliminary analysis on vMOS data; this analysis has been ignored in our previous work; (2) we propose a novel analytical framework in this paper, which is significantly different from the previous work; and (3) experiment results have shown that our proposed model can improve the predication accuracy than our previous work.

In addition, this paper has the following research contributions in contrast to other existing studies: (i) our model consists of K-means clustering approach and logistic regression; the combination of these two approaches can greatly improve the predication accuracy; (ii) we have conducted extensive experiments on three training datasets and one testing dataset and the experiment results show that our proposed model outperforms other existing methods in terms of predication accuracy; and (iii) moreover, our results also imply that a small set of QoS parameters play an important role in determining vMOS.

The remainder of this paper is organized as follows. Section 2 describes the data used in this paper and identifies the challenges. We then present the overview of our method in Section 3. Section 4 presents the experimental results. Finally, we conclude this paper in Section 6.

2 Data description

We obtained the realistic datasets from SpeedVideo Global Operating Platform (SVGOP) established by Huawei[2]; SVGOP is a specific application of vMOS in mobile networks throughout the world. In particular, three datasets contain totally 89,266 samples with 11 features (i.e., QoS parameters) and 1 scoring factor vMOS. Table 2 summarizes the features. As shown in [16], vMOS is a function of the above QoS parameters. However, *to the best of our knowledge, there is no data-driven analysis on the relation between vMOS and other QoS parameters.*

We first conduct a preliminary statistic analysis on the dataset. In particular, Fig. 1 displays the histogram of vMOS in different scales (ranging from 1 to 5), where the vertical axis represents the frequency of vMOS and the horizontal axis represents the evaluation scale of vMOS. We observe from Fig. 1 that most of vMOS values concentrate on the range from 3.5 to 4. When plotting the normal distribution curve (i.e., the red curve in Fig. 1), we can find that the median is about 4.0, implying that most of vMOS values are quite close to "good".

Table 1 Mean Opinion Score (MOS) evaluation scale

Score	Quality	Distortion	Class
1	*Bad*	Very annoying but objectionable	0
2	*Poor*	Annoying, but not objectionable	
3	*Fair*	Perceptible and slightly annoying	1
4	*Good*	Just perceptible, but not annoying	
5	*Excellent*	Imperceptible	

Wang *et al. EURASIP Journal on Wireless Communications and Networking* (2018) 2018:173

Page 3 of 10

Table 2 Description of dataset

Types	Features
QoS parameters	Average rate of playing phase (kbps)
	Video total download (DL) rate (kbps)
	Video bitrate (kbps)
	Initial max DL rate (kbps)
	End-to-End (E2E) round-trip time (RTT) (ms)
	Initial buffering latency (ms)
	Video Initial buffer downloaded (byte)
	Playing time (ms)
	Playing total duration
	Stalling times
	Stalling ratio
Scoring factors	vMOS

We then investigate the correlations between vMOS and other QoS parameters. In particular, Fig. 2 shows a histogram of vMOS versus video total DL rate. It is shown in Fig. 2 that there is a positive correlation between vMOS and video total DL rate, implying that the higher the video total DL rate is, the higher the vMOS is.

In order to measure the correlation between vMOS and initial buffering latency, we plot vMOS against initial buffering latency in Fig. 3 where there is a negative correlation between vMOS and initial buffering latency. In other words, vMOS decreases with the increment of initial buffering latency, implying the worse QoE to users.

Figure 4 plots vMOS against stalling ratio. Essentially, stalling ratio is defined as Stalling ratio $= \frac{\text{Stalling duration}}{\text{Playing total duration}}$ according to the white paper of Huawei [16]. Since most of Stalling ratio values concentrate on
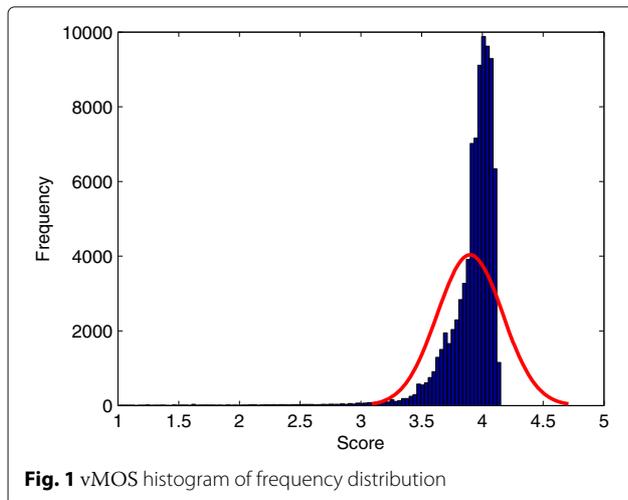
Fig. 1 vMOS histogram of frequency distribution

some small values (i.e., less than 0.3), we also plot a subgraph in Fig. 4 to better present the results. We observe from Fig. 4 that there is a negative correlation between vMOS and stalling ratio. In other words, the higher the stalling ratio is, the lower the vMOS is.

Due to the space limitation, we do not show other correlation analytical results on other QoS parameters in this section. In summary, we find that the dataset has the following characteristics:

- *Heterogeneous data types.* The preliminary results show that QoS parameters are in different types and different ranges. For example, the initial buffering latency is ranging from 500 ms (millisecond) to 30,000 ms while the average rate of play phase is ranging from 300 kbps (kilobit per second) to 16,000 kbps.
- *Positive/negative correlations* of QoS parameters. As shown in the statistics results, we observe that there are positive or negative correlations between vMOS and other QoS parameters. For example, there is a positive correlation between vMOS and video total DL rate while there is a negative correlation between vMOS and initial buffering latency.
- *Dependence* on QoS parameters. The preliminary statistics results also show that vMOS is essentially affected by multiple factors, such as average rate of playing phase, video total DL rate, end-to-end (E2E) round-trip time, and initial buffering latency.

The above characteristics result in the difficulties in analyzing video QoE data. To address the above challenges, we propose a novel data-driven QoE analysis framework (which will be described in details in Section 3).

3 QoE analysis framework

In order to address the above concerns, we propose a novel data-driven QoE analysis method, which consists of the following procedures: (1) data preprocessing and (2) data analysis with integration of K-means clustering and logistic regression. Figure 5 shows the flow chart of our proposed method.

3.1 Data preprocessing

We need to normalize the features since they are in different units. In particular, we make a conversion from the absolute value to the relative value. In particular, we choose the MAX-MIN scaling method to normalize the positive and negative values. More specifically, we have

- Positive values:

$$u_{ij} = \frac{x_{ij} - \min\left(x_{ij}\right)}{\max\left(x_{ij}\right) - \min\left(x_{ij}\right)}, \tag{1}$$

- Negative values:

Wang *et al. EURASIP Journal on Wireless Communications and Networking* (2018) 2018:173

Page 4 of 10

Fig. 2 Histogram of vMOS versus video total DL rate
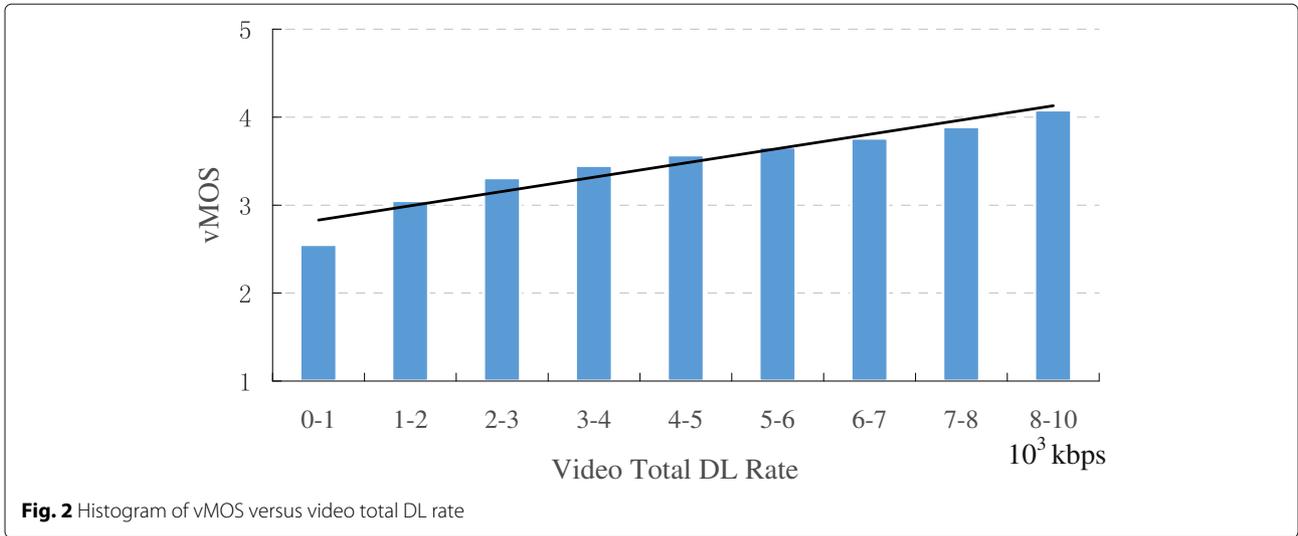
$$u_{ij} = \frac{\max\left(x_{ij}\right) - x_{ij}}{\max\left(x_{ij}\right) - \min\left(x_{ij}\right)}, \qquad (2)$$

where x_{ij} represents the original value, u_{ij} represents the value after normalization, $\min(\cdot)$ is the minimum value, and $\max(\cdot)$ is the maximum value.

3.2 K-means + LR method

In this paper, we use logistic regression (LR) mainly to predict whether a user's QoE of video streaming is "good" or "bad" since it is hard to determine it based on a continuous vMOS value within [1, 5]. However, logistic regression requires that the dependent variable is dichotomou

(binary). Therefore, we first exploit K-means to categorize the sample datasets into two groups according to "good" or "bad". We then use logistic regression to predict the QoE of video streaming.

3.2.1 K-means clustering

The main idea of K-means algorithm [21] is to find a partition such that squared error between the empirical mean of a cluster and the points in the cluster is minimized. In particular, given a dataset $X = \{x_i\}$, $i = 1, 2, ..., m$, we partition them into K disjoint clusters so that the sum of the intra-cluster variances is minimized. We denote the K disjoint clusters by $\mathbf{C} = C_1, C_2, ..., C_k$.

Fig. 3 vMOS versus initial buffering latency

Wang *et al. EURASIP Journal on Wireless Communications and Networking* (2018) 2018:173

Page 5 of 10
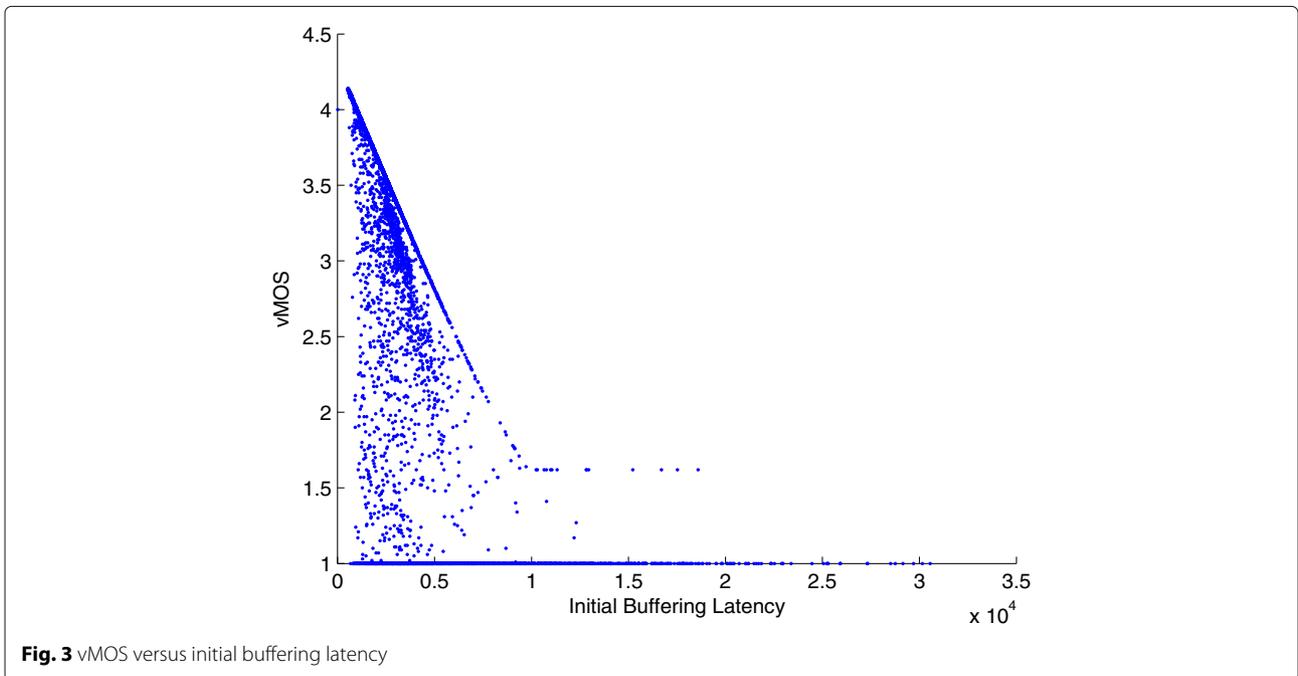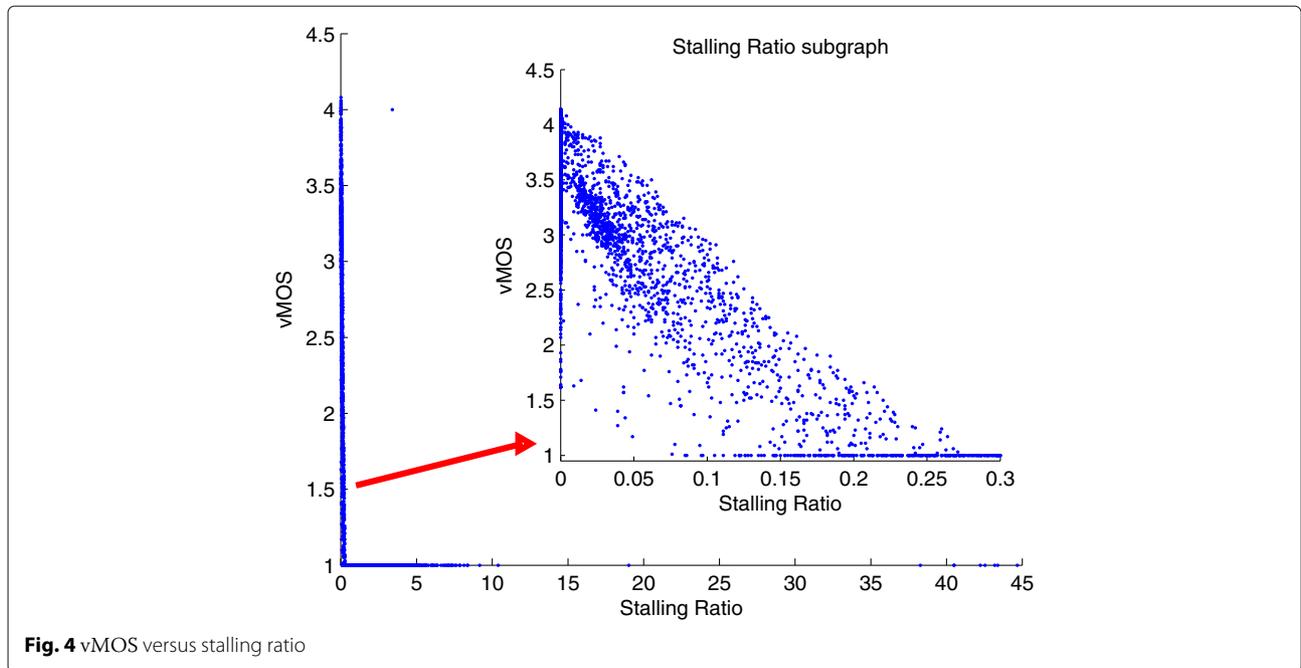
Fig. 4 vMOS versus stalling ratio

The goal of K-means is to minimize the within-cluster sum of variance over all clusters. In other words, we need to find

$$\underset{\mathbf{C}}{\arg\min} \sum_{k=1}^{K} \sum_{x \in C_i} \|x_i - \mu_k\|^2, \tag{3}$$

where μ_k is the mean of cluster C_i.

This clustering process can be completed by alternating between assigning instances to their closest centers and recomputing the centers until a local minimum reaches.

3.2.2 Logistic regression

In this paper, we concern with a binary classification problem of categorizing QoE of a video streaming into two cases of "bad" and "good". To solve this problem, we exploit logistic regression (LR). In particular, we denote the QoE by a binary-dependent variable. This variable only takes two values, either "0" or "1". Specifically, we use "1" to represent "good" and "0" to represent "bad". Without loss of generality, we classify a data sample as "bad" when its vMOS score is within $[1, 2]$ and "good" when its vMOS score is within $[3, 5]$, as shown in Table 1.

Recall that a dataset contains m samples, $D = (X_1, y_1), (X_2, y_2), ..., (X_m, y_m)$, where $X_i = (x_{i1}, x_{i2}, ..., x_{ij})$ is the $i-$th input pattern with dimensionality j and y_i is a corresponding variable that takes a value of 0 or 1. The term $y_i = 0$ indicates the $i-$th sample is bad and $y_i = 1$ indicates the $i-$th sample is good. The vector X_i contains j influence features (for all n QoE features) for the $i-$th sample and x_{ij} denotes the value of feature j for the $i-$th sample. We denote the probability of vMOS being "good" by p and by $1 - p$ of vMOS being "bad". Then, the logit transform of probability p is as follow,

$$\text{logit}(p) = \ln\left[\frac{p}{1-p}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n, \tag{4}$$
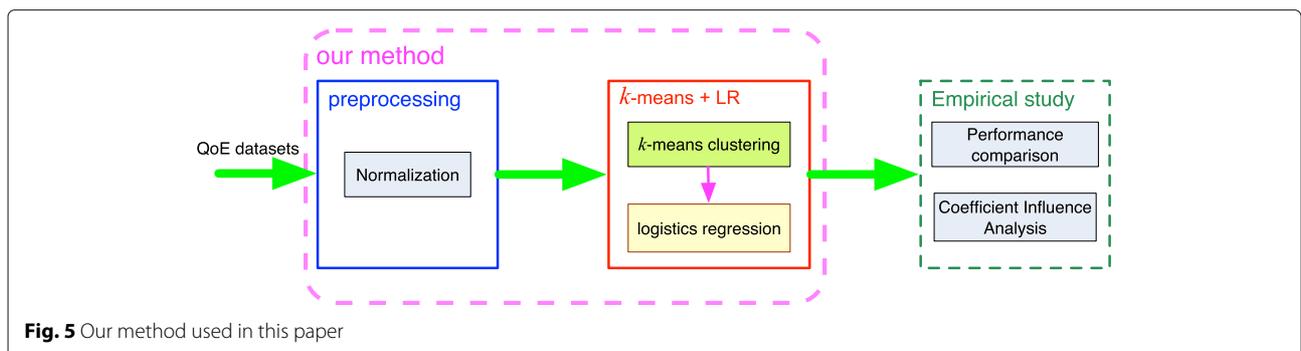
Fig. 5 Our method used in this paper

where β_0 is the offset and β_i $(i = 1, ..., n)$ is the corresponding regression coefficient for each QoS parameter.

It is shown in Eq. (4) that the probability of occurrence denoted by $p(x)$ can be expressed as a non-linear function of features,

$$p(x) = \frac{1}{1 + e^{-\text{logit}(p(x))}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n)}},$$

$$(5)$$

where $p \in [0, 1]$. Since the logistic regression model is non-linear, the maximum likelihood estimation method can be used to estimate the regression coefficient β_i $(i = 0, ..., n)$.

4 Empirical study

As we summarize in Section 2, the dataset has the characteristics such as different types, positive and negative correlations, and dependence. Therefore, we need to use the proposed method to address these concerns. In particular, we describe the experiment settings in Section 4.1. We then show the intermediate results of K-means clustering and logistic regression in Section 4.2. We next compare our proposed method with other existing methods in terms of predication accuracy in Section 4.3. Finally, we conduct performance analysis of our proposed method in Section 4.4.

4.1 Experiment settings

We obtain three sample datasets from SpeedVideo Global Operating Platform of Huawei. Table 3 presents the meta data information of these datasets. In particular, dataset 1, dataset 2, and dataset 3 contain 30,000, 30,000, and 26,984 samples, respectively, each of which has 11 QoS parameters. Table 4 summarizes the 11 features where we denote each of these 11 features by variable x_i $(i = 1$ to $11)$.

4.2 *K*-means LR method

4.2.1 *K-means Analysis*

As shown in Section 3, we exploit the proposed K-means logistic regression method to predict QoE. In particular, we classify each dataset into two classes: one class contains samples with QoE being "good" and another class contains samples with QoE being "bad". The experimental results show that these two classes are fairly unbalanced in terms of class size. Thus, we choose K-means scheme

Table 3 Meta data information

Description	No. of samples	No. of features
Dataset 1	30,000	11
Dataset 2	30,000	11
Dataset 3	26,984	11

Table 4 Summary of features

Types	Features	Variables
QoS parameters	Average rate of playing phase (kbps)	x_1
	Video total download (DL) rate (kbps)	x_2
	Video bitrate (kbps)	x_3
	Initial max DL rate (kbps)	x_4
	End-to-End (E2E) round-trip time (RTT) (ms)	x_5
	Initial buffering latency (ms)	x_6
	Video Initial buffer downloaded (byte)	x_7
	Playing time(ms)	x_8
	Playing total duration	x_9
	Stalling times	x_{10}
	Stalling ratio	x_{11}

to search the cut-off values. Specifically, we obtain classification standard values, 3.9, 3.9, and 3.93 for datasets 1, 2, and 3, respectively. Table 5 shows the K-means clustering results.

4.2.2 *Logistics regression analysis*

We next conduct logistics regression analysis. In particular, we give the regression expression on 11 features as follows,

$$\text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6$$
$$+ \beta_7 x_7 + \beta_8 x_8 + \beta_9 x_9 + \beta_{10} x_{10} + \beta_{11} x_{11}, \quad (6)$$

where x_i $(i = 1$ to $11)$ corresponds to each of 11 features as given in Table 4 and β_i is the regression coefficient.

Table 6 lists regression results on the three datasets, where "Value" denotes the resulting value of each coefficient, "Wal" denotes the value of the Wald test under the significant level of "Sig" [22]. We observe from Table 6 that the lower "Sig" value (or the higher "Wal" value) of a coefficient indicates the higher impact of that coefficient on vMOS. In general, Sig > 0.05, the coefficient is not statistically significant, implying that no correlation can be found between the QoS parameter and vMOS. In practice, there is no correlation between a QoS parameter and vMOS when Sig < 0.05 according to the Wald test [23].

Table 5 *K*-means clustering analysis

Dataset	Classification standard	QoE	No. of samples
Dataset 1	3.90	Bad	10,306
		Good	19,694
Dataset 2	3.90	Bad	19,833
		Good	10,167
Dataset 3	3.93	Bad	9,199
		Good	17,785

Wang *et al. EURASIP Journal on Wireless Communications and Networking*   (2018) 2018:173

Page 7 of 10

Table 6 Regression coefficients

Coefficient	Dataset 1			Dataset 2			Dataset 3		
	Value	Wal	Sig	Value	Wal	Sig	Value	Wal	Sig
β_1	34.43	0.005	0.943	− 15.51	0.272	0.602	− 11.094	0.214	0.643
β_2	− 12.65	0.001	0.98	9.88	0.103	0.748	12.21	0.245	0.621
β_3	105.18	7.374	0.007	1.721	38.976	0	2.263	75.108	0
β_4	− 18.403	0.27	0.599	− 0.632	0.129	0.72	− 0.736	0.152	0.697
β_5	15.89	0.41	0.522	0.313	0.074	0.785	0.943	0.585	0.444
β_6	112529	7.47	0.006	2156.26	404.857	0	2598.501	408.67	0
β_7	− 14.49	0.013	0.909	− 4.778	0.312	0.576	4.1	0.454	0.501
β_8	211762	2.37	0.123	3612.752	6.299	0.012	4204.75	5.961	0.015
β_9	− 3252.45	2.359	0.125	− 56.298	6.409	0.011	− 64.055	5.813	0.016
β_{10}	− 1911.44	0.265	0.607	19.609	0.062	0.803	− 45.572	4.02	0.045
β_{11}	− 207352	2.369	0.124	− 3553.314	6.368	0.012	− 4158.129	4.132	0.042

It is shown in Table 6 that coefficient β_6 (corresponding to initial buffering latency) has the most significant impact on vMOS. Similarly, we can find that coefficient β_3 (corresponding to video bitrate) also has a strong influence on vMOS because of high Wal value (i.e., 7.374).

4.3 Performance comparison

We next present the experimental results over the given datasets with performance comparison with conventional traditional methods including multivariate linear regression, logistic regression, support vector machines, K-nearest-neighbor and Naive Bayes [24].

We then evaluate the performance of our proposed method and other traditional methods. Table 7 presents performance comparison of our proposed K-means + LR method with other methods, where we choose the *precision* as the performance metrics (the higher precision implying the better performance). Among all the methods, our proposed K-means + LR method has the highest performance with precision 96.94, 97.13, and 97.54% on dataset 1, dataset 2, and dataset 3, respectively. This performance improvement may lie in the integration of K-means method and logistic regression method.

4.4 Performance analysis of our method
4.4.1 Residual plots

We first use the residual error plots to evaluate the accuracy of the proposed model. In particular, we define Error = Prediction − Original, where Prediction is the prediction value and Original is the given value. Figure 6 shows the residual plots over the three datasets, in which blue points denote the "good" QoE values and blue points denote the "bad" QoE values. It is shown in Fig. 6 that there are more red points than blue points, implying that our model has a bias of regarding "good" QoE values as "'bad".

4.4.2 Accuracy analysis

In order to assess the accuracy of prediction, we use *Accuracy*, which is defined as follows,

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + FP + TN + FN)}, \qquad (7)$$

where TP is true positive, FP is false positive, FN is false negative, and TN is true negative.

We obtain another dataset from SVGOP. This dataset contains 2517 samples. This new dataset serves as the testing dataset. We next evaluate the accuracy of our proposed model over two datasets: training dataset (i.e., dataset 1) and testing dataset (i.e., the dataset containing 2517 samples). Figure 7 gives the histogram of the accuracy values of 11 QoS parameters over two aforementioned datasets. We observe from Fig. 7 that training dataset has the accuracy value close to that of testing dataset in most of QoS parameters while there are some gaps between several QoS parameters (e.g., playing total duration); the performance gaps may owe to the fact of the small sample size of testing dataset (2517 samples versus 30,000 samples in training dataset). Note that we have similar findings in dataset 2 and dataset 3. Without repetition, we do not show the results on dataset 2 and dataset 3 here.

We next use the *deviance* to evaluate the measure of the fitness of the data. In particular, the deviance denoted by D value is calculated by comparing a given model with the saturated model [24]. We define D value = |prediction − original|, where prediction and original denote the prediction value and the original value, respectively. Table 8 gives the deviance results over the three datasets and the corresponding rankings (sorting D value in ascending order). We observe from Table 8 that initial buffering latency, stalling times, and stalling ratio have the smaller D values

Wang *et al. EURASIP Journal on Wireless Communications and Networking* (2018) 2018:173

Page 8 of 10

Table 7 Performance comparison

Method	Precision		
	Dataset 1	Dataset 2	Dataset 3
Multivariate linear regression	72.55%	72.51%	71.51%
Logistics regression	74.93%	77.55%	83.55%
SVM	82.56%	87.17%	85.42%
K-nearest-neighbor	52.48%	61.34%	56.77%
Naive Bayes	88.12%	83.47%	91.09%
***K*-means + LR (proposed)**	***96.94%***	***97.13%***	***97.54%***

Data in bold is the proposed method that outperforms other existing methods

than other QoS parameters, implying that they essentially have the dominant influence on vMOS. This result is consistent with the previous finding in [20].

In summary, we can see from Table 8 that initial buffering latency, stalling times, and stalling ratio can significantly affect vMOS. In particular, initial buffering latency, stalling times, and stalling ratio are negatively correlated with vMOS. We have the following major findings: (1) *vMOS is affected by multiple QoS parameters together.* Essentially, vMOS is affected by 11 QoS parameters as given in Table 4. (2) *Small set of QoS parameters dominates the performance of vMOS.* Interestingly, we observe that a small set of QoS parameters has the stronger influence on vMOS than other QoS parameters.

5 Discussion and future work
The analytical results essentially offer us some useful insights in improving video QoE, which may pave the way toward improving vMOS in mobile networks. We summarize several possible future directions as follows.

- *Concentrating on optimizing several dominating QoS parameters in vMOS.* For example, we may focus on optimizing the network resource to reduce initial buffering latency, stalling times, and stalling ratio so that we can significantly improve the video QoE while maintaining relatively low operating expenditure. However, it is not an easy task to achieve this goal because the enhancement of these QoS parameters is also involved with many other technologies, such as cross-layer optimization and distributed resource allocation [10, 25, 26].
- *Identifying QoS bottlenecks.* Determining QoS bottlenecks can help to enhance system performance and consequently improving vMOS. However, it is also difficult to identify the QoS bottlenecks since they are often affected by many factors. For example, video stalling is essentially caused by many factors, such as network congestion, network failure, device mobility, and radio spectrum scarcity. There is a challenge in identifying the causality of stalling. In the future, we may apply data-driven approach to identify the reason behind video stalling according to different scenarios.
- *Distributing the videos appropriately to improve vMOS.* For example, we can distribute the most popular videos at the servers close to users so that we can significantly reduce the initial buffering latency.
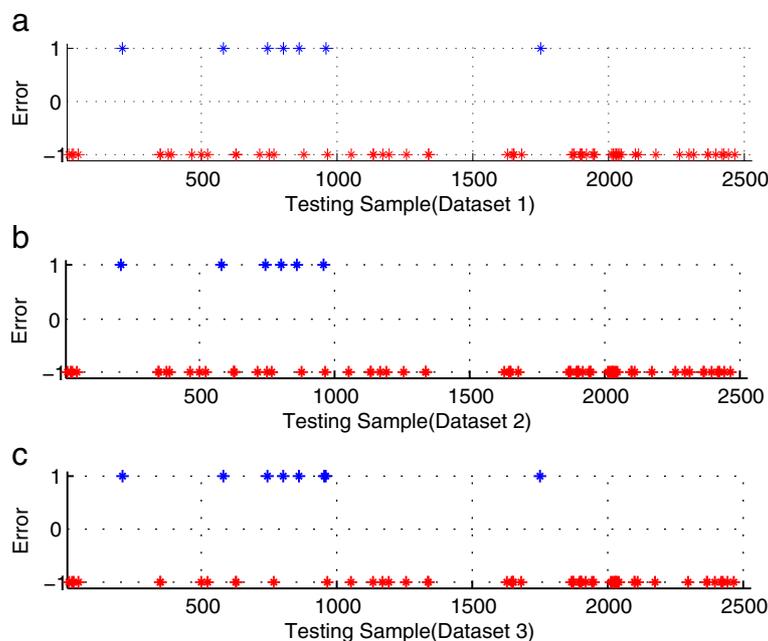
Fig. 6 Residual error plot. **a** Dataset 1. **b** Dataset 2. **c** Dataset 3

Wang *et al. EURASIP Journal on Wireless Communications and Networking* (2018) 2018:173

Page 9 of 10
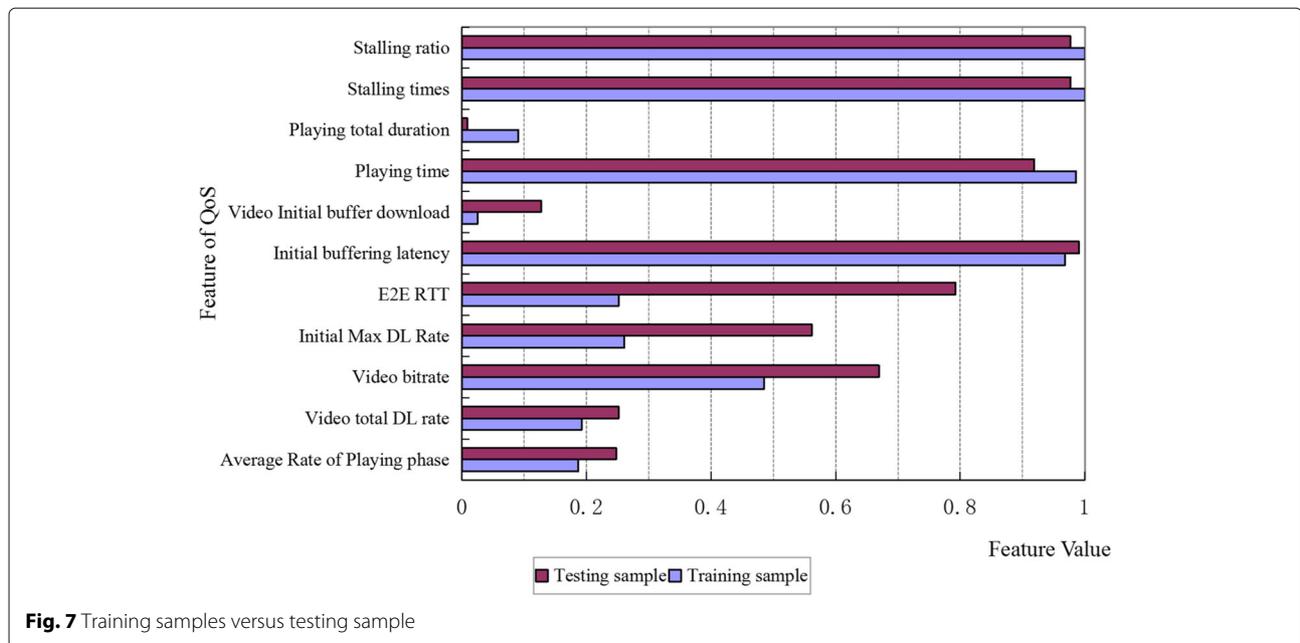
Fig. 7 Training samples versus testing sample

However, to determine the popularity of video streaming is challenging since it requires the extensive efforts in analyzing the massive video data [27].

6 Conclusions

In this paper, we propose a novel data analysis model to analyze video Mean Opinion Score (vMOS), which is an important measure of user quality of experience of video streaming. In particular, our proposed model is a combination of K-mean clustering method and logistic regression method, which can essentially improve the prediction accuracy than other existing methods. We conduct experiments over several realistic datasets. Extensive experiment results show that our proposed method outperforms other existing methods in terms of prediction accuracy. For example, our proposed method has the precision of 96.94, 97.13, and 97.54% on dataset 1, dataset 2, and dataset 3. Our results also show that a small set of QoS parameters play an important role in determining vMOS; this implies that we can concentrate on enhancing these key QoS parameters. It can be achieved by integrating cross-layer optimization and distributed resource allocation schemes together and mitigating QoS bottlenecks.

Our model has a broad range of applications. For example, it can be used to enhance the QoE of video service providers (such as Netflix and YouTube), video-centric mobile applications (including Facebook LIVE, Instagram

Table 8 *D* value and ranking

Features	Dataset 1		Dataset 2		Dataset 3	
	D value	Ranking	*D* value	Ranking	*D* value	Ranking
Average rate of playing phase	0.0605	7	0.0662	7	0.0667	6
Video total DL rate	0.0596	8	0.0651	8	0.0653	8
Video bitrate	0.1848	3	0.1620	3	0.2160	3
Initial max DL rate	0.3008	2	0.2997	2	0.3155	2
E2E RTT	0.5409	1	0.5186	1	0.5602	1
Initial buffering latency	0.0222	11	0.0222	11	0.0199	11
Video initial buffer downloaded	0.1020	4	0.1025	4	0.1033	4
Playing time	− 0.0668	6	− 0.0667	6	− 0.0657	7
Playing total duration	− 0.0814	5	− 0.0741	5	− 0.0695	5
Stalling times	− 0.0231	10	− 0.0229	10	− 0.0279	9
Stalling ratio	− 0.0232	9	− 0.0231	9	− 0.0224	10

Wang *et al. EURASIP Journal on Wireless Communications and Networking* (2018) 2018:173

Page 10 of 10

LIVE, Snapchat, etc.), video game live streaming services (such as Twitch, Hitbox and NetEase Game Lives). Moreover, it can be used to improve the usability of video surveillance systems. For example, the quality of video streaming of video surveillance systems can be helpful in detecting dangers in advance.

Endnotes

[1] For simplicity, we use vMOS to represent U-vMOS throughout this paper.

[2] http://speedvideo.huawei.com/

Abbreviations
DL: Download; E2E: End-to-end; LR: Logistic regression; QoS: Quality of service; QoE: Quality of experience; Sig: Significance; SVGOP: SpeedVideo global operating platform; SVM: Support vector machine; vMOS: video mean opinion score; Wal: Wald test value

Authors' contributions
QW proposed the idea, derived the results, and wrote the paper. H-ND supervised the work and revised versions. DW gave valuable suggestions on the motivation of conducting data analysis on video streaming over mobile networks. HX contributed to motivating, revising and proofreading of the article. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Faculty of Information Technology, Macau University of Science and Technology, Wai Long Avenida, Macau SAR, China. [2]School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510006, China. [3]Guangdong Province Key Laboratory of Big Data Analysis and Processing, Guangzhou 510006, China. [4]Faculty of Computer, Guangdong University of Technology, No. 100 Waihuan Xi Road, Guangzhou 510006, China.

References
1. Cisco, Visual Network Index. "forecast and methodology, 2015-2020". White Paper, 1–41 (2016)
2. M Venkataraman, M Chatterjee, Inferring video QoE in real time. IEEE Netw. **25**(1), 4–13 (2011)
3. Y Chen, K Wu, Q Zhang, From QoS to QoE: A tutorial on video quality assessment. IEEE Commun. Surv. Tutorials. **17**(2), 1126–1165 (2015)
4. H Nam, KH Kim, H Schulzrinne, in *IEEE INFOCOM*. QoE matters more than QoS: Why people stop watching cat videos (IEEE, New Jersey, 2016), pp. 1–9
5. V Joseph, G de Veciana, in *IEEE INFOCOM*. NOVA: QoE-driven optimization of DASH-based video delivery in networks (IEEE, New Jersey, 2014), pp. 82–90
6. A Balachandran, V Sekar, A Akella, S Seshan, I Stoica, H Zhang, in *Proceedings of the ACM SIGCOMM Conference*. Developing a Predictive Model of Quality of Experience for Internet Video (ACM, New York, 2013)
7. S Thakolsri, W Kellerer, E Steinbach, in *Proceedings of IEEE International Conference on Communications (ICC)*. QoE-Based Cross-Layer Optimization of Wireless Video with Unperceivable Temporal Video Quality Fluctuation (IEEE, New Jersey, 2011)
8. V Ramamurthi, O Oyman, J Foerster, in *IEEE Global Communications Conference (GLOBECOM)*. Video-QoE aware resource management at network core (IEEE, New Jersey, 2014), pp. 1418–1423
9. M Venkataraman, M Chatterjee, Effects of internet path selection on video-QoE: analysis and improvements. IEEE/ACM Trans. Networking (TON). **22**(3), 689–702 (2014)
10. J Jiang, V Sekar, H Milner, D Shepherd, I Stoica, H Zhang, in *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*. CFA: A Practical Prediction System for Video QoE Optimization, (Berkeley, 2016), pp. 137–150
11. Y Sun, X Yin, J Jiang, V Sekar, F Lin, N Wang, et al, in *Proceedings of ACM SIGCOMM*. CS2P: Improving video bitrate selection and adaptation with data-driven throughput prediction (ACM, New York, 2016), pp. 272–285
12. W Huang, Y Zhou, Xie X, D Wu, M Chen, E Ngai, Buffer State is Enough: Simplifying the Design of QoE-Aware HTTP Adaptive Video Streaming. IEEE Trans. Broadcast. **64**(2), 1–12 (2018)
13. C Lal, V Laxmi, MS Gaur, M Conti, Enhancing QoE for video streaming in MANETs via multi-constraint routing. Wirel. Netw. **24**(1), 235–256 (2018)
14. V Burger, T Zinner, L Dinh-Xuan, F Wamser, P Tran-Gia, A Generic Approach to Video Buffer Modeling Using Discrete-Time Analysis. ACM Trans Multimedia Comput Commun Appl. **14**(2s), 33:1–33:23 (2018)
15. OE Marai, T Taleb, M Menacer, M Koudil, On Improving Video Streaming Efficiency, Fairness, Stability, and Convergence Time Through Client-Server Cooperation. IEEE Trans. Broadcast. **64**(1), 11–25 (2018)
16. Huawei, "Requirements of Mobile Bearer Network for Mobile Video Service". White Paper, 1–8 (2016)
17. International Telecommunication Union (ITU), "Subjective video quality assessment methods for multimedia applications". ITU Recommendation: 1–37
18. W Pan, G Cheng, H Wu, Y Tang, in *IEEE/ACM 24th International Symposium on Quality of Service (IWQoS)*. Towards QoE assessment of encrypted YouTube adaptive video streaming in mobile networks (IEEE, New Jersey, 2016)
19. CM Lentisco, L Bellido, QJCCE Pastor, JLA H, QoE-Based Analysis of DASH Streaming Parameters Over Mobile Broadcast Networks. IEEE Access. **5**, 20684–20694 (2017)
20. Q Wang, HN Dai, H Wang, D Wu, in *The 15th IEEE International Symposium on Parallel and Distributed Processing with Applications*. Data-Driven QoE Analysis on Video Streaming in Mobile Networks (IEEE, New Jersey, 2017)
21. X Wu, V Kumar, JR Quinlan, J Ghosh, Q Yang, H Motoda, et al, Top 10 algorithms in data mining. Knowl. Inf. Syst. **14**(1), 1–37 (2008)
22. J Friedman, T Hastie, R Tibshirani, The elements of statistical learning. vol. 1. Springer series in statistics New York (2001)
23. RR Wilcox, *Introduction to robust estimation and hypothesis testing, Academic press*. (Elsevier, MA, USA, 2011)
24. E Alpaydin, *Introduction to machine learning*. (MIT press, MA, USA, 2014)
25. Y Wang, X Lin, User-provided networking for QoE provisioning in mobile networks. IEEE Wirel. Commun. **22**(4), 26–33 (2015)
26. Z Su, Q Xu, Q Qi, Big data in mobile social networks: a QoE-oriented framework. IEEE Netw. **30**(1), 52–57 (2016)
27. Z Su, Q Xu, F Hou, Q Yang, Q Qi, Edge Caching for Layered Video Contents in Mobile Social Networks. IEEE Trans. Multimed. **19**(10), 2210–2221 (2017)