End-to-end Quality Adaptation Scheme Based on QoE Prediction for Video Streaming Service in LTE Networks

Huifang Chen[1,2]   Xin Yu[1]   Lei Xie[1,2]

[1]Dept. of Information Science and Electronic Engineering, Zhejiang University
[2]Zhejiang Provincial Key Laboratory of Information Network Technology
No. 38, Zheda Road, Hangzhou 310027, P. R. China
E-mail: { chenhf; xin_yu; xiel}@zju.edu.cn

Abstract—**How to measure the user's feeling about mobile video service and to improve the quality of experience (QoE), has become a concern of network operators and service providers. In this paper, we first investigate the QoE evaluation method for video streaming over Long-Term Evolution (LTE) networks, and propose an end-to-end video quality prediction model based on the gradient boosting machine. In the proposed QoE prediction model, cross-layer parameters extracted from the network layer, the application layer, video content and user equipment are taken into account. Validation results show that our proposed model outperforms ITU-T G.1070 model with a smaller root mean squared error and a higher Pearson correlation coefficient. Second, a window-based bit rate adaptation scheme, which is implemented in the video streaming server, is proposed to improve the quality of video streaming service in LTE networks. In the proposed scheme, the encoding bit rate is adjusted according to two control parameters, the value of predicted QoE and the feedback congestion state of the network. Simulation results show that our proposed end-to-end quality adaptation scheme efficiently improves user-perceived quality compared to the scenarios with fixed bit rates.**

Keywords-Quality of experience (QoE); Gradient boosting machine; Video quality evaluation; Bit rate adaptation

I. INTRODUCTION

Long-Term Evolution (LTE) is a standard developed by the Third Generation Partnership Project (3GPP) for wireless communication of high-speed data service, with the potential to provide increased peak data rates of 100Mbps downstream and 50Mbps upstream [1]. However, due to the difficulties in maintaining high reliability and latency requirements of mobile video applications, it is still a challenging issue to guarantee good user-perceived quality enough for video streaming service over LTE networks.

As we know, traditional quality of service (QoS) can only measure objective network indicators, such as packet loss rate, delay, delay jitter, and so on. However, the quality of experience (QoE), defined as "*the overall acceptability of an application or service, as perceived by the end-user*" by ITU-T [2], not only includes QoS, but also considers the capability of user equipment (UE) as well as user's expectation and context. Hence, QoE is a comprehensive indicator to measure the performance of end-to-end systems, as shown in Fig. 1.

The most direct way to measure QoE for mobile videos is to gather the mean opinion score (MOS), ranging from 1 to 5,

rated by viewers through subjective experiments [3]. However, the process is time-consuming and needs a large amount of manpower, which is practical only if the aim is to benchmark the performance of objective methods.
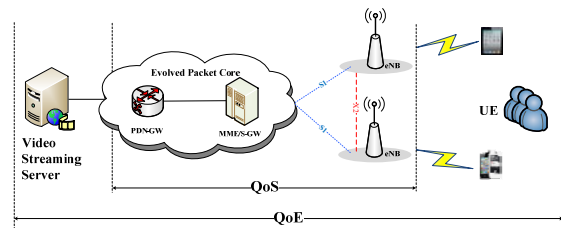

Figure 1.   QoE and QoS for video streaming service in LTE networks

Objective video quality assessment can be divided into full-reference (FR), reduced-reference (RR) and no-reference (NR) methods. Peak-signal-to-noise-ratio (PSNR) is a common FR metric by making a pixel-by-pixel comparison between the source video and degraded video without considering what they actually represent [4]. However, both FR and RR methods require reference information from source videos, which is not available in reality. Hence, only NR method is a feasible way to measure QoE in real-time by modeling the features of transmitted videos.

In [5], the video quality with different content types is measured by multiple linear regression models, in which only the impact of application-level QoS is considered. However, the relationship between QoS and QoE is usually more complicated than the linear ones [6]. In [7], a QoE space with N reference points was proposed. The predicted value equals to the reference point with the nearest Euclidean distance. Since the computational complexity of this method is $O(N^2)$, it is not suitable for a large scale data set. In [8], the concept of feed-forward neural networks is used to predict video quality. However, the video content features are not considered in this method, which have a significant effect on the prediction result of QoE [9]. In addition, an opinion model for video telephone is standardized as G.1070 by ITU-T [10], in which the degradation introduced by network transmission and encoding compression is taken into account. Therefore, to design a good QoE prediction (estimation) model for the mobile video service in wireless networks is very important and urgent, and we will investigate how to predict the QoE for the video streaming service in wireless networks.

On the other hand, the purpose of the research on QoE is to optimize the user experience through video quality adaptation schemes based on predicted value of QoE. A quality adaptation controller was proposed for encoding bit rate switching using feedback control theory in [11], and an adaptive rate control scheme using wireless channel status was developed in [12]. However, these quality adaption schemes did not take the QoE into consideration. Although a novel QoE-driven adaptation scheme at the pre-encoding stage with fuzzy logic technique was proposed in [13], this scheme requires a high computing capability of UE, and to customize the QoE estimation model on UE is not easy to implement. Therefore, we will also focus on investigating the quality adaption scheme based on proposed QoE prediction model.

In this paper, we first present an end-to-end QoE prediction model for video streaming service in LTE networks. In our proposed QoE prediction model, the gradient boosting decision tree (GBDT) algorithm with an ensemble of M base learners is adopted, and cross-layer parameters extracted from the video content, application layer, network layer and UE, are used. Then, a bit rate adaptation scheme implemented at the video streaming server is proposed based on the value of estimated QoE and the feedback information of the network congestion. In order to avoid the bit rate fluctuation, a window-based adjustment strategy is presented. In addition, the system can be implemented for real-time monitoring.

II. GBDT-BASED QOE PREDICTION MODEL

In this section, the QoE prediction model based on GBDT algorithm is presented.

A. Parameters selection and extraction

Video streams suffer from packet loss, delay, delay jitter, and so on when transmitted through the network. Principal component analysis (PCA) proves that packet loss rate (PLR) is the most important QoE parameter in the network layer [13].

Before transmitted, source videos are encoded in the application layer. Encoding parameters, such as bit rate (BR) and frame rate (FR), determine the degree of compression.

The capability of UE contributes to QoE. We assume that intelligent terminals have sufficient capability for video decoding, and tablet PCs (like iPad) require a higher video resolution than mobile phones (like iPhone) due to the larger screen size. Hence, the screen size (SS) of UE and the video resolution (Res) should also be taken into consideration.

In addition, video content features (CF) also affect QoE. For example, the video clips with slight movements require lower bit rate and frame rate, and have a stronger robustness to packet loss compared to the video clips with rapid movements. Our QoE prediction model considers three content-aware parameters derived in temporal and spatial domain.

The temporal information (TI) is defined as the difference between pixels (at the same position) in consecutive frames. More rapid motion in videos will result in larger value of TI [3]. The temporal information can be calculated as

$$M_n(i,j) = F_n(i,j) - F_{n-1}(i,j), \qquad (1)$$

$$TI = \max_{\text{time}}\{std_{\text{space}}[M_n(i,j)]\}, \qquad (2)$$

where $F_n(i,j)$ is the pixel value of the nth frame at the ith row and the jth column, TI is the maximum standard deviation of $M_n(i,j)$.

The spatial information (SI) consisting of the edge block (EB) and the average luminance difference (B_r) is defined to measure the spatial complexity.

The edge block (EB) is extracted with the Sobel filter, where each frame is filtered with the Sobel operator. That is,

$$EB = \max_{\text{time}}\{std_{\text{space}}[Sobel(F_n)]\}. \qquad (3)$$

The average luminance difference (B_r) is defined as the sum of absolute difference of average luminance values between consecutive frames. That is,

$$B_r = \sum_{i=1}^{N}\sum_{j=1}^{M}|\bar{B}_n(i,j) - \bar{B}_{n-1}(i,j)|, \qquad (4)$$

where $\bar{B}_n$ is the average brightness of the nth frame with $N{\times}M$ pixels.

Since the GBDT-based prediction model is implemented at the streaming server, the encoding parameters at the application layer, BR, FR, Res, can be directly obtained from the encoder.

The video content features are calculated from source videos and a content features database can be constructed.

The UE information, SS, is extracted from the international mobile equipment identity (IMEI) when users request video service.

The network-layer parameters are transmitted to the server through RTCP receiver report [21] from UE continuously, which collects QoS information, such as transmitted bytes, packet loss rate, delay, and so on, to present the network congestion level. Hence, pakcet loss rate can be extracted through the "PLR Extractor" from RTCP feedback.

Therefore, all of the parameters needed by the QoE prediction model are available.

B. QoE prediction model

In regression and classification area, a single model is prone to be highly complicated and over-fitted, which can be resolved effectively by a powerful prediction model assembling a lot of base learners. Hence, the gradient boosting machine is proposed. Our proposed QoE prediction model for video streaming service is based on the GBDT machine invented by J.H. Friedman [14].

Let the input vector, $\mathbf{x} = \{x_1, x_2, ..., x_n\}$, be n QoE parameters, $\{PLR, BR, FR, Res, SS, TI, EB, B_r\}$ described in the section II.A, and n=8. The output value is y, corresponding to the predicted MOS value ranging from 1 to 5.

The objective is to find the optimal prediction function, $F^*(\mathbf{x})$, which minimizes the expected value of the loss function, $L(y, F(\mathbf{x}))$, in the training sets. That is,

$$F^*(\mathbf{x}) = \arg\min_{F(\mathbf{x})} E_{y,\mathbf{x}}\left[L(y, F(\mathbf{x}))\right], \qquad (5)$$

where the loss function, $L(y, F(\mathbf{x}))$, uses the least-squared error (LSE). That is,

$$L(y, F(\mathbf{x})) = (y - F(\mathbf{x}))^2. \qquad (6)$$

The prediction function $F(\mathbf{x})$ can be expressed by a set of parameters, $\mathbf{P} = \{P_1, P_2, ...\}$, and the additive expression is given as

$$\tilde{F}(\mathbf{x}; \mathbf{P}) = \sum_{m=1}^{M} \beta_m h(\mathbf{x}; \boldsymbol{\alpha}_m), \qquad (7)$$

where $\mathbf{P} = \{(\beta_m, \boldsymbol{\alpha}_m) \mid m = 1, 2, ..., M\}$, and $h(\mathbf{x}; \boldsymbol{\alpha}_m)$ is the form of the mth base learner.

For the training set $\{(\mathbf{x}_i, y_i) \mid i = 1, 2, ..., N\}$ with N points, the optimization problem formulated in (5) can be rewritten as

$$(\beta_m, \boldsymbol{\alpha}_m) = \arg\min_{\beta, \boldsymbol{\alpha}} \sum_{i=1}^{N} L(y_i, F_{m-1}(\mathbf{x}_i) + \beta h(\mathbf{x}_i; \boldsymbol{\alpha})),$$
$$m = 1, 2, ..., M \qquad (8)$$

The optimization problem formulated in (8) can be resolved with an iteration process as follows.

(1). *Initialization*

Defining an initial base learner with a constant value, ρ, as

$$F_0(\mathbf{x}) = \arg\min_{\rho} \sum_{i=1}^{N} L(y_i, \rho). \qquad (9)$$

(2). *Base learner construction*

A new base learner, $h(\mathbf{x}; \boldsymbol{\alpha}_m)$, is trained on the residuals of outcomes from all previous learners in each iteration process. In order to make the reduction of the residual steepest, the base learner is constructed in the "steepest-descent" direction, $-g_m(\mathbf{x})$, which is an inverse direction of the gradient of $L(y, F(\mathbf{x}))$. That is,

$$-g_m(\mathbf{x}) = -\left[\frac{\partial L(y_i, F(\mathbf{x}_i))}{\partial F(\mathbf{x}_i)}\right]_{F(\mathbf{x}) = F_{m-1}(\mathbf{x})}, i = 1, 2, ..., N. \quad (10)$$

$\boldsymbol{\alpha}_m$ is the vector that makes $h(\mathbf{x}; \boldsymbol{\alpha}_m)$ approach along $-g_m(\mathbf{x})$,

$$\boldsymbol{\alpha}_m = \arg\min_{\boldsymbol{\alpha}} \sum_{i=1}^{N} \left[-g_m(\mathbf{x}_i) - \beta_m h(\mathbf{x}_i; \boldsymbol{\alpha})\right]^2. \qquad (11)$$

With the steepest-descent direction, β_m is given by the line search as

$$\beta_m = \arg\min_{\beta} \sum_{i=1}^{N} L(y_i, F_{m-1}(\mathbf{x}_i) + \beta h(\mathbf{x}_i; \boldsymbol{\alpha}_m)). \qquad (12)$$

(3). *Updating the prediction function*

Finally, the base learner in the mth iteration is obtained as

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \beta_m h(\mathbf{x}; \boldsymbol{\alpha}_m). \qquad (13)$$

To avoid over-fitting, the base learner is usually multiplied by a learning rate, v, ranging from 0 to 1. That is,

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + v \cdot \beta_m h(\mathbf{x}; \boldsymbol{\alpha}_m). \qquad (14)$$

In (14), a small v means more iteration times, which results in more computation time, and $v \to 1$ will lead to over-fitting in a faster learning process. The common value of v is around 0.1, which is a trade-off between the prediction accuracy and the computational complexity.

Based on theoretical analysis mentioned above, a QoE prediction model for video streaming service is set up as illustrated in Fig. 2. The base learners (BLs) are in the form of regression trees. The input parameters of the GBDT-based QoE prediction model, $\{PLR, BR, FR, Res, SS, TI, EB, B_r\}$, are end-to-end packet loss rate from the network layer, the video bit rate, frame rate and resolution from the application layer, the screen size from UE, and the content features from videos. More special, video content features are derived through the "feature extraction" module.

In the GBDT-based QoE prediction model, there are two processes, namely training process and predicting process, represented by the dashed line and solid line in Fig. 2, respectively. In the training process, objective cross-layer parameters in combination with subjective MOS values are used to train M base learners with good generalization ability. In the predicting process, only objective parameters are available, and the predicted QoE values are obtained from the well-trained model.
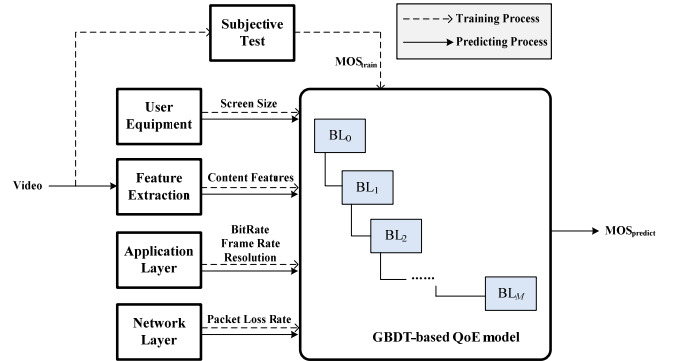

Figure 2. The GBDT-based QoE prediction model

C. Validation of the proposed model

For validation, we choose 6 source video clips ranging from slight movements to rapid movements in database libraries [15] and [16], including *akiyo*, *mother-daughter*, *foreman*, *city*, *soccer* and *ice*. These raw videos are in YUV (4:2:0) format with a length of 8~12s.

First, they are encoded with H.264/AVC (MPEG-4 Part 10) format, where the group of pictures (GOP) pattern IBBPBBPBB is adopted. The encoded H.264 bit-streams have different resolutions (QCIF and CIF), bit rates (32kbps~1Mbps) and frame rates (10, 15, and 30fps).

Second, these H.264 bit-streams are transmitted through LTE network simulated under NS2 platform as shown in Fig. 3. The evaluation model consists of a video streaming server, a router, aGW standing for network elements in evolved packet core, an eNodeB (eNB) and a fixed UE. UE connects to eNB through wireless link with DLAirQueue for downlink and ULAirQueue for uplink, which are implemented to simulate the air interface in LTE [17].
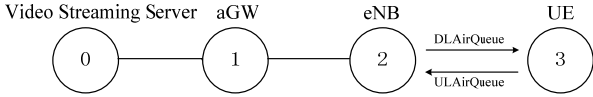

Figure 3.   LTE network topology

The simulated transmission process is based on the Evalvid module modified by C.-H. Ke for NS2 [18]. The radio channel is the connection bottleneck with burst packet loss rate modeled with 2-state Markov process [19]. The simulation parameters are given in Table I.

TABLE I.   SIMULATION PARAMETERS

| Parameters	Value
Bandwidth and delay (server-aGW)	100Mbps, 10ms
Bandwidth and delay (aGW-eNB)	10Mbps, 2ms
Downlink bit rate and delay	2Mbps, 5ms
Uplink bit rate and delay	512Kbps, 10ms
Maximum packet size	1000bytes
UDP header size	8bytes
IP header size	20bytes
Packet loss rate	0~20%

After transmission, the degraded videos are available at the UE side. We emulate a larger screen size (iPad) and a smaller screen size (iPhone) with changeable video playback windows. Finally, we get about 800 decoded videos with varying degree of damage. Since there is a lack of testers to rate such a large dataset, video quality is measured by average PSNR and converts to MOS value from Evalvid-RA [20].

From the degraded videos, 80% data are randomly chosen to train the GBDT-based QoE prediction model, and the rest data are used to evaluate the performance of the model. We construct 100 base learners to make a 4-fold cross validation and compare the prediction results with the G.1070 QoE model, the results are shown in Fig. 4.

From Fig. 4, we observe that the points of GBDT-based QoE model are centered around the diagonal line, which means that predicted MOS values are close to objective MOS values. This benefits from comprehensive input parameters and assembled regression tree models. Since G.1070 QoE model only considers the impact of packet loss rate and encoding compression, the predicted MOS values give a larger deviation to objective MOS values. Another reason is that G.1070 QoE model is designed for video telephone service, and is not suitable for video streaming service well.
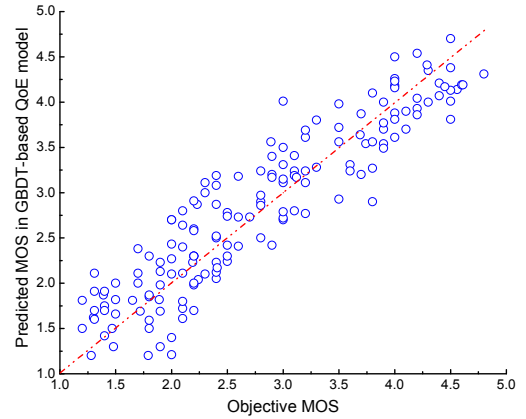
(a) GBDT-based QoE prediction model
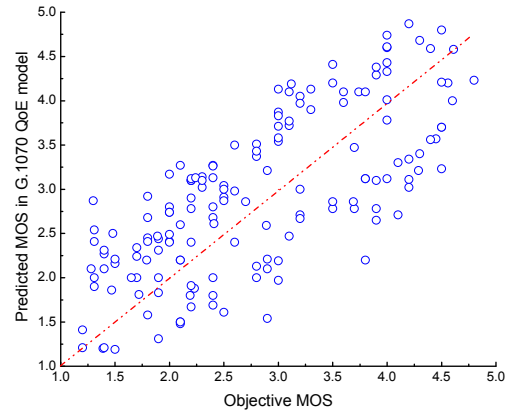
(b) G.1070 QoE prediction model

Figure 4.   The performance comparision of two QoE prediction models

Table II shows the performance comparison between two models, in terms of the root mean squared error (RMSE) and Pearson correlation coefficient. From Table II, we can see that the prediction result of our proposed GBDT-based QoE model is more accurate with a lower RMSE and a higher Pearson correlation coefficient.

TABLE II.   COMPARISON OF TWO QOE PREDICTION MODELS

Model	RMSE	Pearson Coefficient
GBDT-QoE	0.323468	0.829127
G.1070	0.673824	0.693691

III.   QOE-DRIVEN QUALITY ADAPTATION SCHEME

In this section, we present an application of the GBDT-based QoE prediction model in a window-based bit rate adaptation scheme implemented at the video streaming server.

A.   System model

As we know, compared with other services, such as voice, best effort service and IPTV, video streams increase the data traffic significantly when transmitted in wireless networks. If the load of traffic is high enough, the network will become congested and some packets will be discarded, and transmission delay will also increase. In this case, the video quality continues to deteriorate and the QoE will become increasingly worse without congestion control. Hence, a

quality adaptation mechanism is necessary to help control the congestion and maintain an adequate QoE value.

Here, we propose a bit rate adaptation scheme based on estimated QoE value and the feedback congestion information from UE to improve the user experience, which is also implemented at the video streaming server. The system architecture is illustrated in Fig. 5.
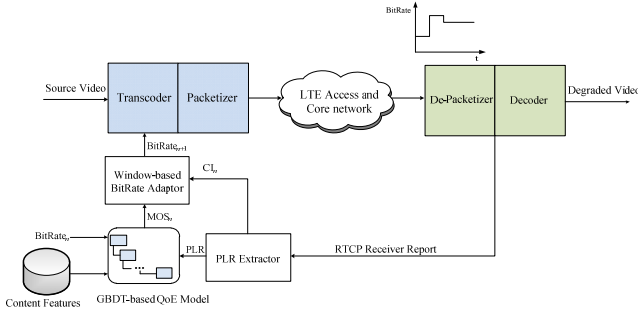

Figure 5.   The system architecture

By measuring QoE value at the server, the system monitors user experience more efficiently and reliefs the computation burden from UE compared to the method in [13].

B.   Bit rate adaption strategy

Since we mainly focus on bit rate adaptation strategy in this work, it is assumed that the video resolution and the screen size of UE is adapted well, and the video frame rate is fixed. Hence, in the view of bit rate adapation, the reasons for QoE degradation are mainly two aspects. First, the encoding bit rate is too low so that wireless resource is wasted if the channel condition is good. Second, the network is under congestion, and video packets are discarded from the interface queue, so the video quality will continue to deteriorate if the encoding bit rate is not adjusted.

The proposed bit rate adaptor, as shown in Fig. 5, is controlled by two parameters, namely the estimated MOS value and the congestion indicator (CI) during the nth period. The congestion indicator is measured by PLR, where CI=0 means that the network is in a good state ($PLR \leq 0.1\%$) and CI=1 means that the network is in a congestion state ($PLR > 0.1\%$). The output of the bit rate adaptor is the encoding bit rate for the next period, the input of a transcoder.

The encoding bit rate is partitioned into four levels, $level$=0, 1, 2 and 3 corresponding to low, medium, high and excellent bit rates, respectively. In the first adaptation period, the initial bit rate is medium in order to get a small access delay. Then, at the beginning of the $(n+1)$th transcoding period, the estimated QoE value during the nth period, MOS_n, is obtained from the GBDT-based QoE prediction model. If $MOS_n \geq MOS_{th}$, no adapation is performed. Since frequent adaptation is annoying to viewers, MOS_{th} can be set to 3.5 according to [13]. If $MOS_n < MOS_{th}$, which means user experience is under a satisfied level, the bit rate adaptation should be taken.

To avoid bit rate fluctuation, we adopt a sliding window mechanism with a length of two periods to observe the congestion indicators, CI_{n-1} and CI_n, during the $(n-1)$th period and the nth period. Because the state of the wireless channel is time-varying, long observation time is not necessary. Here, there are four cases, where $level_{n-1}$, $level_n$ and $level_{n+1}$ represent the bit rate class during the $(n-1)$th, nth, $(n+1)$th period.

Case 1: $CI_{n-1} = 0$ and $CI_n = 0$. It indicates that the network is in a good state during the past two periods. In this case, since QoE degradation results from insufficient bit rate, the adjusted bit rate is $level_{n+1} = \min\{\min\{level_{n-1}, level_n\} + 1, 4\}$.

Case 2: $CI_{n-1} = 0$ and $CI_n = 1$. It indicates that the network condition begins to decrease. In this case, the bit rate should be reduced to alleviate the congestion, and then $level_{n+1} = \max\{\min\{level_{n-1}, level_n\} - 1, 0\}$.

Case 3: $CI_{n-1} = 1$ and $CI_n = 0$. It indicates that the network has been recovered from a congestion state. Hence, the bit rate will not to be changed, and $level_{n+1} = level_n$.

Case 4: $CI_{n-1} = 1$ and $CI_n = 1$. It indicates the network is in a serious congestion. Hence, the bit rate should be reduced significantly, and $level_{n+1} = \max\{\min\{level_{n-1}, level_n\} - 2, 0\}$.

The principle of the bit rate adaptation scheme is to gradually increase the bit rate when the network is in a good state, and quickly reduce the bit rate when the network is in a congestion state. Based on our proposed bit rate adaptation scheme, a timely bit rate adjustment can be executed at the video streaming server when the QoE prediction model reports an unsatisfied MOS value.
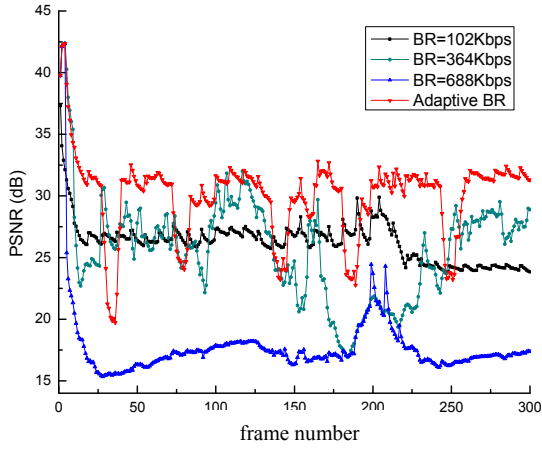
C.   Performance evaluation and analysis

In this subsection, we evaluate the performance of the proposed bit rate adaptation scheme. The *foreman* video with a length of 300 frames is chosen since it includes both fast movements and slight movements. Then, the source video is encoded with a CIF resolution and a constant frame rate (30fps). The transmission topology is shown in Fig. 3, and the downlink bandwidth is changeable in order to emulate different wireless link conditions. The bit rate switching is simulated in Evalvid-RA, and the adaptation period is set to be a GOP size. Four encoding levels are 94Kbps, 255Kbps, 516Kbps and 912Kbps, representing the low, medium, high and excellent class in bit rate adaptation scheme, respectively.

Fig. 6 shows the PSNR of every frame encoded with three constant bit rates and an adaptive bit rate with different downlink bandwidth, where wireless downlink bandwidth is 512kbps, 768kbps and 1500kbps for Figs. (a), (b) and (c), respectively. From these figures, we observe that, when wireless downlink bandwidth is small, the PSNR of video encoded with low constant bit rate, such as BR=102Kbps, is more stable than that of encoded with high constant bit rates and adaptive bit rate. When wireless downlink bandwidth increases, the PSNR of video encoded with high constant bit rates, such as BR=364Kbps and BR=688Kbps, will become stable. The reason for this phenomenon is that, when the
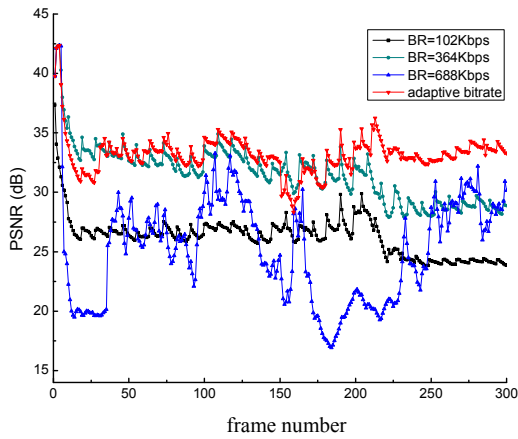
wireless downlink bandwidth is small, video encoded with lower constant bit rate will not result in network congestion, while video encoded with higher constant bit rate, *BR*=688Kbps, will result in serious packet loss and lead to the worst performance due to a large data traffic. However, for

(a) wireless link bandwidth=512Kbps

(b) wireless link bandwidth=768Kbps

(c) wireless link bandwidth=1500Kbps

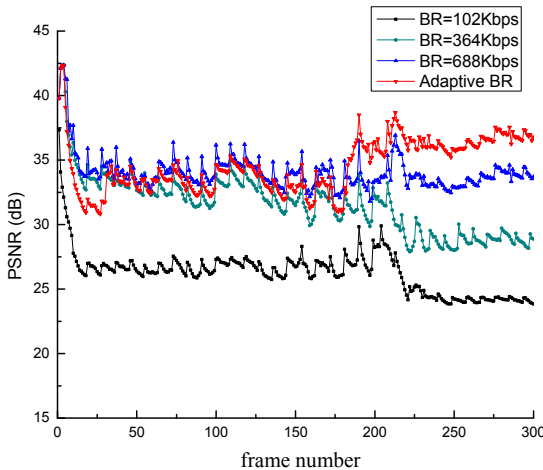Figure 6. The PSNR of every frame with constant bit rates and an adaptive bit rate

our proposed bit rate adaptation scheme, the video quality always remains at a relatively high average PSNR due to effective QoE prediction model and the bit rate adjustment strategy although the bit-stream is initialized with medium encoding bit rate.

Fig. 7 shows the same reconstructed frame encoded with different bit rates, where the wireless downlink bandwidth is 768kbps. From Fig. 7, the reconstructed frame in Fig. 7(a) is blurry due to a low bit rate of 102Kbps, and the reconstructed frame in Fig. 7(c) with a high bit rate of 688Kbps is disturbed by the packet loss. The reconstructed frames in Figs. 7(b) and 7(d) remain at an acceptable level encoded with 364Kbps and an adaptive bit rate with PSNR of 33.45dB and 34.53dB, respectively.

(a) *BR*=102Kbps

(b) *BR*=364Kbps

(c) *BR*=688Kbps

(d) adaptive bit rate

Figure 7. The same reconstructed frame with different bit rates

Fig. 8 shows the comparison of average objective MOS between constant bit rates and adaptive bit rate when wireless downlink bandwidth is 512Kbps, 768Kbps and 1.5Mbps, respectively. From Fig. 8, we observe that average objective MOS of video encoded with adaptive bit rate is larger than that of encoded with constant bit rates for three different conditions of wireless downlink bandwidth. And the bit rate adaptation scheme obtains objective MOS of 3.02 under a low bandwidth and 4.22 under a high bandwidth. The reason for this phenomenon is that the encoding bit rate is adjusted to avoid or alleviate network congestion, and achieve higher bandwidth utilization. Hence, QoE is improved.

IV. CONCLUSIONS

QoE is a comprehensive indicator to measure the performance of end-to-end systems. In this paper, we investigate the QoE prediction model and the video quality improvement scheme based on the proposed QoE prediction model.
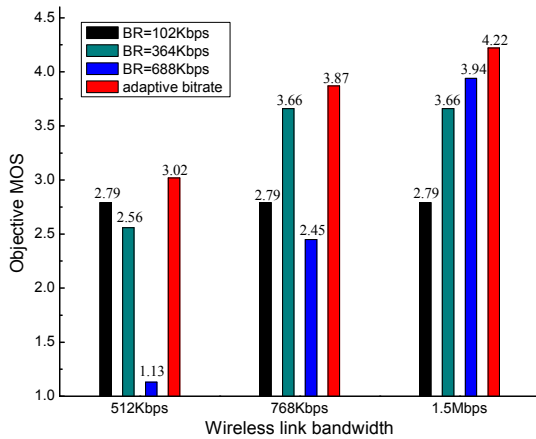
Figure 8. The comparison of objective MOS under different values of wireless downlink bandwidth

The contributions are twofold. First, an end-to-end QoE prediction model based on GBDT machine is proposed for video streaming service in LTE networks. In proposed QoE prediction model, cross-layer parameters affecting the value of predicted QoE are taken into consideration. Performance evaluation results show that the prediction performance of our proposed GBDT-based QoE model outperforms the G.1070 QoE model with a smaller RMSE value and a higher Pearson correlation coefficient. Second, we develop a QoE-driven bit rate adaptation system to improve user experience. The bit rate adaptation scheme is based on the estimated QoE value and the feedback network congestion state. Simulation results show that our proposed bit rate adaptation scheme, consisting of a bit rate adaptation strategy and a sliding window mechanism with a length of two periods to observe the congestion indicator, improves the user experience efficiently compared to the scheme with constant bit rates.

REFERENCES

[1] 3GPP TS 136.300, LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 10), version 10.7.0, Mar. 2012.

[2] "Definition of Quality of Experience", Ref: TD 109rev2 (PLEN/12), ITU-International Telecommunication Union, Jan. 2007.

[3] "Subjective video quality assessment methods for multimedia applications", ITU-T Recommendation P.910, 2008.

[4] S. Winkler, P. Mohandas, "The evolution of video quality measurement: from PSNR to hybrid metrics", IEEE Trans. on Broadcasting, vol. 54, pp. 660-668, Sept. 2008.

[5] F. Agboma, A. Liotta, "Quality of experience management in mobile content delivery systems", Journal of Telecommunication Systems, vol. 49, no. 1, pp. 85-98, 2012.

[6] M. Fiedler, T. Hossfeld, T. Phuoc, "A generic quantitative relationship between quality of experience and quality of service", IEEE Journal of Network, vol. 24, pp. 36-41, Mar. 2010.

[7] M. Venkataraman, M. Chatterjee, S. Chattopadhyay, "Evaluating quality of experience for streaming video in real time", Proc. of IEEE GLOBECOM 2009, pp. 1-6, Nov. 2009.

[8] P. Calyam, P. Chandrasekaran, G. Trueb, "Multi-resolution multimedia QoE models for IPTV applications", International Journal of Digital Multimedia Broadcasting, vol. 2012, pp. 136-149, Jun. 2012.

[9] K. Asiya, L. Sun, I. Emmanuel, "Content-based video quality prediction for MPEG-4 video streaming over wireless networks", Journal of Multimedia, vol. 4, no. 4, pp. 228-239, Aug. 2009.

[10] "Opinion model for video-telephony applications", ITU-T Recommendation G.1070, Apr. 2007.

[11] D. C. Luca, M. Saverio, P. Vittorio, "Feedback control for adaptive live video streaming", Proc. of ACM Multimedia Systems 2011, pp. 145-156, 2011.

[12] K. Jahon, C. Kwangsue, "MARC: adaptive rate control scheme for improving the QoE of streaming services in mobile broadband networks", Proc. of IEEE ISCIT 2010, pp. 105-110, Oct. 2010.

[13] A. Khan, I.-H. M. Mkwawa, L. Sun, E. C. Ifeachor, "QoE-driven sender bitrate adaptation scheme for video applications over IP multimedia subsystem", Proc. of IEEE ICC 2011, pp. 1-6, Jun. 2011.

[14] J. H. Friedman, "Greedy function approximation: a gradient boosting machine", Journal of the Annals of Statistics, pp. 1189-1232, 2001.

[15] Standard test sequences, available online at ftp://ftp.tnt.uni-hannover.de/pub/svc/testsequences/.

[16] YUV Sequences, available online at http://trace.eas.asu.edu/yuv.

[17] Q. Qiu, J. Chen, L. Ping, Q. Zhang, X. Pan, "LTE/SAE model and its implementation in NS2", Proc. of IEEE MSN 2009, pp. 299-303, Dec. 2009.

[18] C.-H. Ke, C.-K. Shieh, W.-S. Hwang, A. Ziviani, "An evaluation framework for more realistic simulations of MPEG video transmission", Journal of Information and Science and Engineering, vol. 24, no. 2, pp. 425-440, Mar. 2008.

[19] E. N. Gilbert, "Capacity of a burst-noise channel", Technical Journal of the BELL System, vol. 39, pp. 1253-1265, Sep. 1960.

[20] A. Lie, J. Klaue, "Evalvid-RA: Trace driven simulation of rate adaptive MPEG4 VBR video", Journal of Multimedia System, vol. 14, no. 1, pp/ 33-50, 2008.

[21] T. Freidman, R. Caceres, A. Clark, "RTP control protocol extended reports (RTCP XR)", Nov. 2003.