Cross-network and Cross-layer Optimized Video Streaming over LTE and WCDMA Downlink

Zhenjie Deng*, Yanwei Liu *, Jinxia Liu†, Xin Chen*, Antonios Argyriou§, Zhen Xu* and Song Ci‡
*State Key Laboratory of Information Security,
Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
†Zhejiang Wanli University, Ningbo, China
‡University of Nebraska-Lincoln, Omaha, USA
§University of Thessaly, Volos, Greece
Email: {dengzhenjie,liuyanwei}@iie.ac.cn

***Abstract*—Video services are proliferating over today's mobile Internet and efforts have been made to improve their performance. The cross-layer video streaming that jointly optimizes the parameters at different protocol layers is a feasible solution but without bandwidth aggregation of multiple wireless networks. The cross-network video streaming can realize the bandwidth aggregation by using multiple overlapping wireless networks simultaneously. However, parameters at different protocol layers are optimized independently in the cross-network optimization. In this paper, we propose a joint cross-network and cross-layer optimized video streaming scheme that utilizes the bandwidth aggregation of cross-network video streaming and further improve the performance by jointly optimizing the parameters of different protocol layers in each network with a cross-layer manner. In the proposed scheme, the LTE and WCDMA networks are adopted. The bit-rate of the video at application layer, the rate allocation among networks and the parameters of physical layers in each network are jointly optimized. Experimental results show that the proposed scheme gains higher quality of experience in terms of PSNR than the state-of-the-art schemes.**

Keywords*—video streaming; cross-network optimization; cross-layer optimization; LTE; WCDMA

I. INTRODUCTION

Video streaming over mobile Internet has become exceedingly popular over the past few years. YouTube and Netflix, as two major video content providers, both predict that their respective mobile video traffic will grow explosively within the next few years [1]. Meanwhile, the resolution of video content is rapidly increasing along with the evolution of advanced image capturing technology. For example, videos with high-definition (HD) or even higher spatial resolution are widespread over the Internet. However, due to the bandwidth limitations or the unpredictable and unreliable nature of wireless networks, streaming HD video over wireless network nowadays still poses a big challenge.

In particular, nowadays streaming HD video to a mobile user still encounters the following issues [2]. First, mobile users sometimes suffer from insufficient bandwidth, which will result in a long start-up delay, frequent re-buffering and declining playback quality. In addition, as the channel of mobile network is highly time-variant, video streaming without rate adaption will further degrade the performance. Therefore, to efficiently transfer a HD video in wireless network, it is necessary to establish a steady connection with sufficient bandwidth between the client and the video content provider and make the content bit-rate match with the connection.

To achieve the above goals and ensure the user experience of HD video streaming service, researchers have proposed some video streaming optimization techniques in recent years. The common one is that video content is partitioned into segments and each segment is independently encoded at different bit-rates to enable the adaptive video streaming [3]. In addition, some cross-layer approaches that jointly optimize the parameters at different protocol layers were adopted for video streaming [4-8]. However, the performance of HD video streaming over one wireless network is still hindered by its limited bandwidth, which is generally the bottleneck of transmission.

Today, mobile devices are gradually equipped with multiple wireless interfaces connected with different overlapping networks (Wi-Fi or cellular 3G/4G) [2]. And also, the next generation wireless network will be heterogeneous, for instance, LTE cellular network is becoming prevalent while the legacy networks such as WCDMA are still widely used [9]. Since WCDMA and LTE are operating in different frequency ranges, mobile devices with two long-range wireless radio interfaces can potentially access to them simultaneously [9, 20]. In the future wireless networks, this multiple radio access technology (multi-RAT) can be realized and is a viable solution to fulfill the explosive growth of mobile traffic. Therefore, a potential solution to the problem of insufficient bandwidth is the cross-network video transmission that delivers the video data over different wireless networks simultaneously for bandwidth aggregation. There have been a lot of researches [10-14] that attempted to address this cross-network video streaming problem, but most of them focus on how to efficiently utilize multiple wireless networks at the application layer, and the parameters at other protocol layers that can be jointly coordinated for optimization in each network are usually neglected.

To fully take advantage of the cross-network and the cross-

* Yanwei Liu is corresponding author.
This work was supported in part by Strategic Pilot Project of Chinese Academy of Sciences under Grant XDA06010306, Public welfare project of Zhejiang Province under Grant 2014C31072 and Ningbo Natural Science Foundation under Grant 2015A610133.

layer (CNCL) optimized video streaming, a mobile video streaming scenario is considered in this paper that a multi-homed client can request the HD video segments through LTE and WCDMA access networks simultaneously in a cross-network way. In the application layer, the video rate and the rate allocation among the corresponding networks are selected to minimize the expected video distortion. In addition, the parameters in the physical layer such as modulation and channel coding scheme (MCS) in LTE and spreading factor (SF) in WCDMA are tuned to optimize the cross-network HD video streaming with a cross-layer manner. Our main contributions can be summarized as follows:

- Firstly, a cross-network video streaming optimization scheme that jointly exploits the LTE and WCDMA downlink is proposed to provide the function of bandwidth aggregation. The optimal bit-rates and the rate allocation for each connected network are adaptively determined for each streaming segment to cooperate the demand of the video stream with the dynamic wireless channels.
- Secondly, to jointly optimize the bit-rate selection of the video at the application layer and the parameter configurations at the physical layer in each connected network, a cross-layer optimization scheme that utilizes the MCS mode in LTE downlink and the SF in WCDMA downlink is proposed to combine with the cross-network optimization scheme to minimize the expected end-to-end video distortion.

The rest of the paper is organized as follows. Section II presents the related work. The proposed joint CNCL video streaming system model are presented in Section III. The joint CNCL optimization problem formulation as wel as the solution are described in Section IV. In Section V, the experimental results are presented. Finally, Section VI concludes this paper.

II. Related work

The related work to this paper can be generally classified into two categories: cross-layer optimized video streaming and cross-network optimized video streaming.

Cross-layer optimized video streaming is a commonly discussed topic in recent years. Specifically, the parameters at different protocol layers are jointly considered to optimize the video streaming. With respect to CDMA network, Chen et al. [4] proposed a cross-layer optimization scheme called traffic-adaptive scheme for multimedia transport over multi-code CDMA networks. The rate-adaptive scheme could significantly reduce the frame loss ratio, increase the system throughput and optimize the packet delivery delay of the system. The authors in [5] proposed a cross-layer scheme that jointly optimizes the outer loop SINR-target and variable spreading factor at the physical layer to minimize the packet error rate in the CMDA network. With regard to LTE network, Zhao et al. [6] proposed a SSIM-based cross-layer optimized video streaming over LTE downlink. The modulation and coding scheme at the physical layer of the LTE downlink is selected by jointly considering the characteristics of the video packet. In [7], a distortion-fair cross-layer resource allocation scheme was proposed to optimize the scalable video transmission in OFDMA wireless networks. The authors in [8] designed a cross-layer optimization framework for efficient video streaming over LTE networks using scalable video coding (SVC) and concluded that approximately thirteen percent video quality gains were observed for the users at the cell edge.

Besides the cross-layer related video streaming techniques, there have been a range of cross-network optimization schemes for wireless video streaming. Among them, Markov Decision Process (MDP) was commonly used to model the process of cross-network video streaming. Lee et al. [10] proposed a MDP based solution to the cost minimization for video-on-demand services in a heterogeneous wireless networks with multi-homed terminals. The solution included the parameter estimation, threshold adjustment, and threshold compensation. Xing et al. [11] proposed a real-time adaptive algorithm for video streaming over multiple access networks. The optimal video streaming process was also formulated as a MDP and a reward function was designed to take the QoS requirement into account. Using the streaming control transmission protocol (SCTP) at the transport layer, the authors in [12] proposed a distortion-aware concurrent multipath transfer scheme for mobile video streaming to minimize the end-to-end video distortion over heterogeneous wireless networks. At the application layer, Song et al. [13] proposed a probabilistic multipath video streaming scheme, which sent video traffic bursts over multiple available channels according to a probability generation function of packet delay. However, the parameter configurations at the physical layer are not addressed in the above two schemes. With respect to rate allocation in cross-network video streaming, Zhu et al. [14] studied the distributed rate allocation policies for multi-homed video streaming over heterogeneous access networks and showed that media-aware rate allocation outperformed the heuristic AIMD-based schemes in terms of perceived video quality.

Though the aforementioned video streaming schemes can achieve some performance improvements, most of them either stress mainly on cross-network or cross-layer independently, which lead to sub-optimal performance. To fully utilize the advantages of both schemes, we focus on a joint cross-network and cross-layer video streaming optimization scheme, which dynamically allocates appropriate video bit-rates to the LTE network and WCDMA network simultaneously, with jointly optimizing the MCS of LTE network and the SF of WCDMA network to make the video traffic at the application layer efficiently adapt to the time-varying channel state at the physical layer.

III. Joint CNCL video streaming system model

A. Overview of the proposed scheme

Fig. 1 shows the proposed joint CNCL video streaming optimization system. The system consists of a joint CNCL optimization controller which resides in the client, two different kinds of base stations (nodeB and eNodeB) that allocate the corresponding physical layer parameters to the client and a video content server that contains multiple copies of video segments with various bit-rates. Assume that there is a bit-rate set $\mathcal{R} = \{r_1, r_2, \cdots, r_n\}$ of copies of video segments available for the client, where $n \geq 2$. A multi-homed client embedded with the CNCL optimization controller is expected to minimize the end-to-end video distortion for each ongoing segment. To achieve this, while the multi-homed client requests for a new segment, the MCS of LTE downlink and the SF of CDMA downlink at the physical layer and the video bit-rate at the application layer are jointly adjusted by the controller as shown in Fig. 1. Meanwhile, the optimal rate allocation is also properly selected to split the video traffic among the connected network.

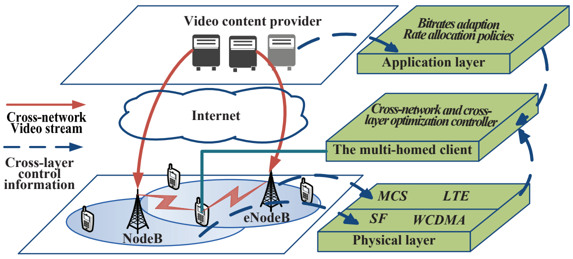

Fig. 1. The joint CNCL video streaming optimization scheme.

B. Video distortion model

In the proposed scheme, a commonly recognized distortion model [15] is utilized to characterize the end-to-end video quality. In the model, the expected end-to-end video distortion at the client consists of the source encoding-induced distortion and the transmission loss-introduced distortion. The former is mostly determined by the distortion-rate (DR) characteristic of encoded video segment. It decays with increasing encoding bit-rate r. The decay is rather steep in the low bit-rate range, but becomes quite low at high bit-rate [12]. The distortion caused by packet loss due to transmission errors or late arrivals is sensitive to the average packet loss probability ρ of all the communication networks. Thus, we can explicitly formulate the expected end-to-end video distortion $E\{d_c\}$ in terms of Mean Square Error (MSE) as

$$E\{d_c\} = d_0 + \frac{\alpha}{r - r_0} + \kappa\rho \tag{1}$$

where d_0 is the source distortion offset, α is the rate-distortion factor, r_0 is the rate factor and κ is the sensitivity factor that reflects the impact of packet loss. All the parameters depend on the encoding structure and the video content, and can be estimated from three or more off-line trial encodings using non-linear regression techniques [12, 15]. They are stored in the server-side and delivered to the controller for optimization during the streaming system set-up.

C. The adjustable physical layer parameters of LTE downlink and WCDMA downlink

1) The capacity and the packet loss probability of LTE downlink with respect to MCS: Since part of the video segment is allocated to the LTE downlink, the capacity and the packet loss probability in LTE downlink will directly affect the performance of CNCL-optimized video streaming system. In order to estimate the packet loss rate for the transmitted video packets, the mutual information effective SINR mapping (MIESM) is exploited to measure the LTE downlink quality in this paper. Assume that there is a set $\mathcal{M} = \{1, 2, \cdots, m\}$ of MCS modes available for transmitting the video packets. For the candidate MCS mode m, the effective SINR mapping $\gamma_{mieff}(m)$ based on the mutual information is computed as [16].

$$\gamma_{mieff}(m) = \kappa(m)[J^{-1}(\frac{1}{N_{sb}}\sum_{j=1}^{N_{sb}} J(\sqrt{\frac{\gamma_j}{\kappa(m)}}))]^2 \tag{2}$$

where N_{sb} is the number of the subcarrier, γ_j is the SINR at the j^{th} subcarrier, and $\kappa(m)$ is the calibration factor for the MCS mode m. The functions of $J(\cdot)$ and $J^{-1}(\cdot)$ are defined as eq. (3) and (4).

Based on the MIESM $\gamma_{mieff}(m)$, the block error rate (BLER) $BLER(\gamma_{mieff}(m))$ for the resource block (RB) with MCS mode m can be accurately formulated as

$$BLER(\gamma_{mieff}(m)) = \frac{1}{2}erfc(\frac{\gamma_{mieff}(m) - b(m)}{\sqrt{2} \cdot c(m)}) \tag{5}$$

where $erfc(\cdot)$ is the complementary error function, $b(m)$ and $c(m)$ are the "transition center" and "transition width" respectively. The values of $b(m)$ and $c(m)$ under different m are shown in Table I, which can be obtained by fitting (5) to the exact BLER in a specific communication system.

In LTE downlink video streaming, one video slice might be carried by several RBs at the physical layer. Therefore, the video packet loss probability $\rho_l(m)$ of a slice is related to the BLERs for all the RBs that the video slice contains as

$$\rho_l(m) = 1 - \prod_{i}^{RB_{num}} (1 - BLER_i(\gamma_{mieff}(m))) \tag{6}$$

where RB_{num} is the RB number that video slice occupies and $BLER_i(\gamma_{mieff}(m))$ is the BLER for the i^{th} RB in the video packet corresponding to the slice.

For LTE system the obtainable bit-rate per symbol rate depends on the SINR. In order to estimate the capacity $\mu_l(m)$ of LTE, the so-called MCS mode index ranging from 1 to 15 is used. By apply the MCS mode in Table I, we get $\mu_l(m)$ as the following [19]:

$$\mu_l(m) = B \cdot r(m) \tag{7}$$

where B denotes the bandwidth of the channel and is 1.4MHz in this paper, and $r(m)$ is the efficient bit-rate per symbol for MCS mode index m (as given by Table I).

2) The capacity and the packet loss probability of WCDMA downlink with respect to SF configuration: WCDMA is a well-known radio communication technique that allows multiple users to share the same wireless spectrum simultaneously. To satisfy different QoS requirements, the WCDMA system has to provide variable data rates. Such flexibility can be achieved by using the orthogonal variable spreading factor (OVSF) codes as the channelization codes [17]. To estimate the packet loss rate for the transmitted video packet, an amended SINR mapping is used to measure the WCDMA downlink channel quality. Assume that the multiple access interference at the receiver obeys the Gaussian distribution [5]. The SINR at the receiver is given by

$$\gamma(f) = \{\frac{k-1}{3f} + \frac{N_0}{2E_b\Omega}\}^{-1} \tag{8}$$

where f is the spreading factor and typically in WCDMA standard $f \in \mathcal{SF} = \{4, 8, 16, 32, 64, 128, 256, 512\}$, E_b/N_0 denotes the energy per bit to noise power spectral density ratio, Ω indicates the path strength and k is the number of users in a given base station. Based on the SINR $\gamma(f)$, the bit error rate (BER) with QPSK modulation can be calculated as

$$BER = Q(\sqrt{\gamma(f)}) \tag{9}$$

$$J(x) \approx \begin{cases} -0.04210610x^3 + 0.209252x^2 - 0.00640081x, & 0 < \mathrm{x} < 1.6363 \\ 1 - \exp(0.00181491x^3 - 0.142675x^2 - 0.08220540x + 0.0548608), & \mathrm{x} \geq 1.6363 \end{cases} \quad (3)$$

$$J^{-1}(y) \approx \begin{cases} 1.09542y^2 + 0.214217y + 2.33727\sqrt{y}, & 0 < \mathrm{y} < 0.3646 \\ -0.706692\log(-0.386013(\mathrm{y}-1)) + 1.75017\mathrm{y}, & 0.3646 \leq \mathrm{y} \leq 1 \end{cases} \quad (4)$$

TABLE I. THE CANDIDATE LTE DOWNLINK MCS MODES

MCS mode(m)	Modulation order	Rate(bits /symbol $r(m)$)	$\kappa(m)$	$b(m)$	$c(m)$
1	QPSK	0.1523	3.07	-7.758	0.6003
2	QPSK	0.2344	4.41	-5.724	0.5182
3	QPSK	0.3770	0.60	-3.652	0.4032
4	QPSK	0.6010	1.16	-1.593	0.3588
5	QPSK	0.8770	1.06	0.3501	0.2910
6	QPSK	1.1758	1.06	2.348	0.2563
7	16QAM	1.4766	0.87	4.297	0.2563
8	16QAM	1.9141	1.01	6.214	0.2293
9	16QAM	2.4063	1.04	8.242	0.2253
10	64QAM	2.7305	1.03	10.13	0.2248
11	64QAM	3.3223	1.11	12.06	0.2028
12	64QAM	3.9023	1.01	13.89	0.1962
13	64QAM	4.5234	1.07	15.72	0.1958
14	64QAM	5.1152	1.00	17.50	0.2134
15	64QAM	5.5547	1.05	19.59	0.2592

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-u^2/2} du$ is the complementary cumulative distribution function. Assume that block codes with l-bits error correction capability are utilized in this work. The instantaneous packet loss probability $\rho_w(f)$ of the video packet with length L that is allocated to WCDMA downlink can be written as

$$\rho_w(f) = 1 - \sum_{i=0}^{l} \binom{L}{i} BER^i (1 - BER)^i \quad (10)$$

Considering that BER is expected to be very small, the packet loss probability can be approximated by

$$\rho_w(f) \approx \frac{(BER)^{l+1} L!}{(l+1)!(L-l-1)!} = \binom{L}{l+1} (BER)^{l+1} \quad (11)$$

The Shannon capacity is used to estimate the capacity allocated to the client by WCDMA network. Based on the measured SINR, the available capacity can be obtained as

$$\mu_w(f) = \tfrac{w}{f} \log(1 + \gamma(f)) \quad (12)$$

where w is the available chip rate that equals to 3.84 Mchip/s in WCDMA standard. It can be observed that the capacity and the packet loss probability of the WCDMA downlink are related to the selected spreading factor f. In the paper, a 4-subscribers, AWGN CDMA downlink channel with different spreading code configurations is adopted.

IV. JOINT CNCL OPTIMIZATION PROBLEM FORMULATION AND SOLUTION

A. Problem formulation

While requesting a new video segment, the multi-homed client first collects the physical layer information and estimates available capacity and packet loss probability of corresponding networks. Based on the effective bandwidth estimation and the feedbacks from different layers, the controller attempts to minimize the expected video distortion for each segment without causing the *rebuffering events*. The decision variables include the requested segment rate r, the sub-rate r_l and r_w allocated to the LTE and WCDMA downlink respectively, the MCS mode m and the spread factor f at the physical layer. We are now ready to formulate the joint CNCL optimization problem for expected video distortion minimization in the following way:

$$\begin{aligned} & \min_{\{r, r_l, r_w, m, f\}} \quad E\{d_c(r, r_l, r_w, m, f)\} \\ s.t. \quad & r = r_l + r_w \\ & \rho = (\rho_l(m) \cdot r_l + \rho_w(f) \cdot r_w)/r \\ & m \in \mathcal{M} \\ & f \in \mathcal{SF} \\ & r \in \mathcal{R} \\ & \max(\tfrac{r_l}{\mu_l(m)}, \tfrac{r_w}{\mu_w(f)}) < T_{\max} \end{aligned} \quad (13)$$

Note that the requested segment is logically separated into two subsegments, which can be implemented by *HTTP's range retrieval requests* [2]. One of the video subsegments is delivered by the LTE network while the remaining subsegment is carried by the WCDMA network. Thus, the requested segment rate r is the summation of two sub-segment rate r_l and r_w. To estimate the overall packet loss probability ρ of the ongoing segment, it can be statistically measured by the second constraint of (13) for simplicity. To avoid the rebuffering events that strongly impair the user's experience, the current requesting segment has to arrive at the client before the playback buffer goes empty. Let T_{max} be the total play time of the downloaded yet unviewed segment remained in buffer. The value of T_{max} can be obtained by periodical feedbacks from the client to the optimization controller. In the joint CNCL video streaming, rebuffering events occurs if one of the sub-segments cannot arrive at the client before the playback buffer runs out. In other words, the rebuffering events will not emerge if $\max(\frac{r_l}{\mu_l(m)}, \frac{r_w}{\mu_w(f)}) < T_{\max}$, which is the last constraint in (13).

B. The proposed solution

An enumerating method or exhaustive search method can be used to solve problem (13), which requires a considerable computational complexity and is unacceptable. To solve the joint CNCL optimization problem efficiently, we construct a heuristic algorithm to find the optimal decision variables $(r^*, r_l^*, r_w^*, m^*, f^*)$ to minimize the expected distortion of the requested segment. That is $E\{d_c(r^*, r_l^*, r_w^*, m^*, f^*)\} \geq E\{d_c(r, r_l, r_w, m, f)\}, \forall (r, r_l, r_w, m, f)$ subject to the constraints defined in (13).

Since the video distortion decreases with increasing bitrate r and decreasing packet loss probability ρ, In the proposed solution, the client aggressively requests the new segment

Algorithm 1: The heuristic joint CNCL optimization algorithm

Input: $\mathcal{R}, \mathcal{M}, \mathcal{SF}, T_{max}, \gamma$
Output: r^*, r_l^*, r_w^*, m^*, f^*
Initialization: $\mathcal{R}_e \leftarrow \mathcal{R}$
while $\mathcal{R}_e \neq \emptyset$ **do**
  $r^* = \arg\max_{r \in \mathcal{R}_e} r$
  while $\mathcal{M} \neq \emptyset, \mathcal{SF} \neq \emptyset$ **do**
    $m^* = \arg\min_{m \in \mathcal{M}} m$
    $f^* = \arg\max_{f \in \mathcal{SF}} f$
    estimate $\rho_l(m^*)$, $\rho_w(f^*)$, $\mu_l(m^*)$ and $\mu_w(f^*)$ according to γ
    $r_l^* = \frac{\mu_l(m^*)}{\mu_l(m^*)+\mu_w(f^*)} \cdot r^*$
    $r_w^* = \frac{\mu_w(f^*)}{\mu_l(m^*)+\mu_w(f^*)} \cdot r^*$
    if $\max(\frac{r_l^*}{\mu_l(m^*)}, \frac{r_w^*}{\mu_w(f^*)}) < T_{\max}$ **then**
      return $(r^*, r_l^*, r_w^*, m^*, f^*)$
    end
    $\mathcal{M} \leftarrow \mathcal{M} \setminus m^*$
    $\mathcal{SF} \leftarrow \mathcal{SF} \setminus f^*$
  end
  $\mathcal{R}_e \leftarrow \mathcal{R}_e \setminus r^*$
end

with the highest bitrate values in the candidate set $\mathcal{R}_e$ and meanwhile avoids packet loss events.That is $r^* = \arg\max_{r \in \mathcal{R}_e} r$. After selecting the optimal bitrate r^* at the application layer, the algorithm will determine the MCS mode and the spreading factor for the corresponding networks at the physical layer. Intuitively, the MCS mode with small constellation and powerful channel code, the spreading factor with large value maintain reliability at poor channel condition. Therefore, the MCS mode with the smallest constellation and channel code, and the spreading factor with the largest value are selected, which are $m^* = \arg\min_{m \in \mathcal{M}} m$ and $f^* = \arg\max_{f \in \mathcal{SF}} f$ in Algorithm 1. Based on the optimal decision variables m^* and f^*, the achievable bandwidth $\mu_l(m^*)$ and $\mu_w(f^*)$ are estimated.

The rate allocation that determines the size of the subsegment transported by each network is based on its achievable throughput under a specific MCS mode or spread factor. That is performed by $r_l^* = \frac{\mu_l(m^*)}{\mu_l(m^*)+\mu_w(f^*)} \cdot r^*$ and $r_w^* = \frac{\mu_w(f^*)}{\mu_l(m^*)+\mu_w(f^*)} \cdot r^*$. Then the optimal decision variables $(r^*, r_l^*, r_w^*, m^*, f^*)$ are obtained so far. However, such a decision variable set might lead to rebuffering event and every subsegment has to arrive at the client before the received buffer runs out. Note that if $\max(\frac{r_l^*}{\mu_l(m^*)}, \frac{r_w^*}{\mu_w(f^*)}) < T_{\max}$, the rebuffering event will not occur and the decision variables are verified and returned. Otherwise, given the selected MCS mode and spreading factor, the achievable bandwidth cannot satisfy the quality level of video segment with bitrate of r^*. As a result, the MCS mode with larger constellation size and more powerful channel code, the smaller spreading factor are selected. The summary of the proposed heuristic algorithm is shown in Algorithm 1.

V. Experimental Results

Based on the H.264/AVC reference software JM16.1, we implement the CNCL-optimized video streaming scheme. The LTE downlink and the WCDMA downlink are simulated by matlab based on [5] and [18], respectively. The specific experimental parameters are shown in Table II. To evaluate the performance of the proposed CNCL optimization scheme, the state-of-the-art schemes including the cross-network (CN) without cross-layer(CL) [12] and and the cross-layer (CL) without CN (without adaptive rate allocation) [6] are used as the baseline schemes. These three schemes are tested at different channel states (standard normal distribution with average SINR at 2dB, 4dB, 9dB, 14dB, 20dB).

TABLE II. EXPERIMENTAL PARAMETERS

Coding profile	Baseline
Frame rate	25fps
Error concealment	Temporal replacement
Coding structure	IPPP
Candidate MCS mode	Table I
Candidate SF	8-256(2^k)
Video resolution	720P
Average SINR	(2,4,9,14,20)dB
Segment durations	2 seconds

Firstly, whether the proposed scheme can correctly obtain the anticipated results is investigated. Fig. 2 shows the optimal rate allocation for the LTE and WCDMA network for 100 segments under the channel condition of average SINR $\overline{\gamma} = 4dB$ for video sequence *Shield*. From Fig. 2, it can be seen that the optimal rate allocation including r_l^* and r_w^* is dynamically tuned to minimize the expected end-to-end distortion for each ongoing video segment. Therefore, it illustrates that the proposed scheme effectively adjusts the MCS, SF at the physical layer and bit-rates of the video with a cross-layer manner and adaptively allocates the bit-rates of the video among networks for cross-network bandwidth aggregation.

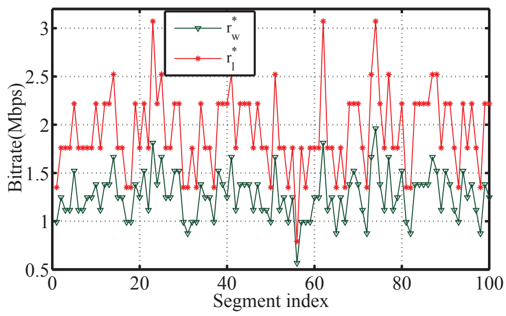

Fig. 2. The optimal bit-rate allocation curves with increasing segment number for the video sequence of *Shield* with the condition of $\overline{\gamma} = 4dB$.

To evaluate the performance of the proposed scheme, the PSNR values for 200 frames of the video sequences *Shield* and *Parkrun* of 720P at the condition of $\overline{\gamma} = 4dB$ are shown in Fig. 3. From Fig. 3, it can be observed that the proposed scheme can achieve higher instantaneous PSNR values than the other two baseline schemes for most of segments. The average PSNR of *Shield* achieved by the proposed scheme is 1.5dB and 2.7dB higher than the two baseline schemes. For the sequence of *Parkrun*, the average PSNR improvement is 2.3dB and 3.5dB, respectively.

Fig. 4 shows the average PSNR curves of the video sequences *Shield* (720P) and *Parkrun* (720P) at different channel conditions. It can be observed from Fig. 4 that the proposed CNCL optimization scheme can achieve higher PSNR values than the other two baseline schemes. On average, the proposed scheme is about 1.74dB and 2.82dB higher than the baseline schemes respectively for the sequence $Parkrun$. For the video sequence of $Shield$, the improvement is approximately 1.24dB and 2.28dB respectively. Furthermore, the performances of average PSNR versus SINR show some differences when SINR falls in different regions. When SINR is with small value, the proposed scheme has a higher improvement in terms of average PSNR than the condition of average SINR with high values. For example, for the sequence of $Shield$ at the condition of $\overline{\gamma} = 2dB$, the average PSNR achieved by the proposed scheme is approximately 2dB and 3.9dB higher than the other two baseline schemes respectively. While at the condition of $\overline{\gamma} = 20dB$, the improvement is just about 0.5dB and 0.8dB, respectively. This is because the proposed scheme can adaptively select the bit-rates of the video, adaptively allocate the video traffic to the LTE network and WCDMA network to utilize their aggregated bandwidth and adjust the parameters of MCS and SF at the physical layers to meet the demand of the video stream and adapt to the fluctuated wireless channels.

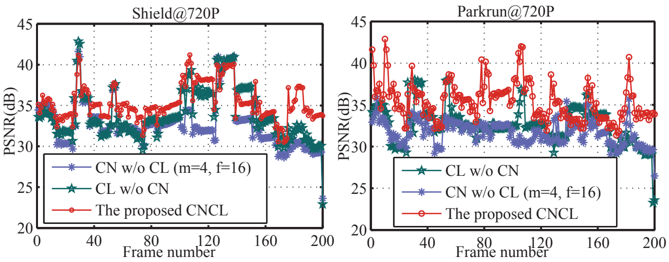

Fig. 3. PSNR curves with increasing frame numbers for the sequence (a) $Shield$ (720P) and (b) $Parkrun$ (720P) at the condition of $\overline{\gamma} = 4dB$.

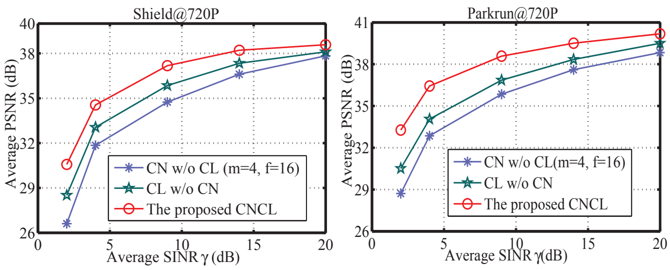

Fig. 4. The average PSNR values of the proposed and the baseline schemes for the sequences (a) $Shield$ (720P) and (b) $Parkrun$ (720P) at different channel conditions.

VI. Conclusion

In this paper, a CNCL optimization scheme is proposed to improve the performance of mobile video streaming. Two wireless access networks, LTE and WCDMA, are used to achieve the bandwidth aggregation of the cross-network scheme. Meanwhile, the cross-layer scheme is combined with the cross-network scheme by optimizing the corresponding parameters (MCS mode and SF) at the physical layer of each network, the bit-rate of the video and the rate allocation at the application layer. Experimental results show that the proposed scheme achieves better streaming performance by improving the average PSNR about 1.49dB and 2.55dB higher than the CL without CN scheme and the CN without CL scheme, respectively.

References

[1] Y. Chen, D. Towsley, and R. Khalili, "Msplayer: Multi-source and multi-path leveraged youtuber," *Proc. 10th ACM Int'l Conf. Emerging Networking Experiments and Technologies (CoNEXT),* pp. 263-270, Dec. 2014.

[2] K. Evensen, T. Kupka, and D. Kaspar, et al., "Quality-adaptive scheduling for live streaming over multiple access networks," *Proc. 20th ACM int'l Workshop on Network and operating systems support for digital audio and video (NOSSDAV),* pp. 21-26, Jan. 2010.

[3] X. Yin, V.Sekar, and B, Sinpoli, "Toward a principled framework to design dynamic adaptive streaming algorithms over HTTP," *Proc. 13th ACM Workshop on Hot Topics in Networks (HotNets),* pp. 9-15, Oct. 2014.

[4] H. Chen, H. C. Chan, and V. Leung, "Cross-layer optimization for multimedia transport over multicode CDMA networks," *IEEE Trans. Mobile Computing,* vol.10, no. 6, pp. 810-820, June. 2011.

[5] A. Shojaeifard, F. Zarringhalam, and M. Shikh-Bahaei, "Packet error rate (PER)-based cross-layer optimization of CDMA networks," *Proc. IEEE Global Telecomm. Conf. (GLOBECOM),* pp. 1-6, Dec. 2011.

[6] P. Zhao, Y. Liu, and J. Liu, et al., "SSIM-based cross-layer optimized video streaming over LTE downlink," *IEEE Global Comm. Conf. (GLOBECOM),* pp.1394-1399, Dec. 2014.

[7] S. Cicalo and V. Tralli, "Distortion-fair cross-layer resource allocation for scalable video transmission in OFDMA wireless networks," *IEEE Trans. Multimedia,* vol. 16, no. 3, pp. 848-863, April. 2014.

[8] R. Radhakrishnan and A. Nayak, "Cross-layer design for efficient video streaming over LTE using scalable video coding," *IEEE Int'l Conf. Comm. (ICC),* pp. 6509-6513, June. 2012.

[9] Y. Mai and J. Chen, "IP multimedia relay architectures with multi-RAT support in LTE-advanced wireless network," *IEEE 7th Asia Modelling Symposium (AMS),* pp. 283-288, July. 2013.

[10] J. Lee and S. Bahk, "On the MDP-based cost minimization for video-on-demand services in a heterogeneous wireless network with multihomed terminals," *IEEE Trans. Mobile Computing,* vol. 12, no. 9, pp. 1737-1749, Sept. 2013.

[11] M. Xing, S. Xiang, and L. Cai, "A real-time adaptive algorithm for video streaming over multiple wireless access networks," *IEEE J. Selected Areas in Comm.,* vol. 32, no. 4, pp. 795-805, April. 2014.

[12] J. Wu, B. Cheng, and C. Yuen, et al., "Distortion-aware concurrent multipath transfer for mobile video streaming in heterogeneous wireless networks," *IEEE Trans. Mobile Computing,* vol. 14, no. 4, pp. 688-701, April. 2015.

[13] W. Song and W Zhuang, "Performance analysis of probabilistic multipath transmission of video streaming traffic over multi-radio wireless devices," *IEEE Trans. Wireless Comm.,* vol. 11, no. 4, pp. 1554-1564, April. 2012.

[14] X. Zhu, P. Agrawal, and J. Singh, et al., "Distributed rate allocation policies for multihomed video streaming over heterogeneous access networks," *IEEE Trans. Multimedia,* vol. 11, no. 4, pp. 752-764, June. 2009.

[15] K. Stuhlmuller, N. Farber, and M. Link, et al., "Analysis of video transmission over lossy channels," *IEEE J. Selected Areas in Comm.,* vol. 18, no. 6, pp. 1012-1032, June. 2000.

[16] T. Jensen, S. Kant, and J. Wehinger, et al., "Fast link adaptation for MIMO-OFDM," *IEEE Trans. Vehicular Technology,* vol. 59, no. 8, pp. 3766-3778, Oct. 2010.

[17] Y. Tseng and C. Chao, "Code placement and replacement strategies for wideband CDMA OVSF code tree management," *IEEE Trans. Mobile Computing,* vol. 1, no. 4, pp. 293-302, Oct-Dec. 2002.

[18] M. Taranetz, T. Blazek, and T. Kropfreiter, et al., "Runtime precoding: enabling multipoint transmission in LTE-advanced system level simulations," *IEEE Access,* vol. 3, no.3, pp. 725-736, June. 2015.

[19] O. Osterbo, "Scheduling and capacity estimation in LTE," *ACM Proc. 23rd International Teletraffic Congress (ITC),* pp. 63-70, Sept. 2011.

[20] B. Bangerter, S. Talwar, R. Arefi and K. Stewart, "Networks and devices for the 5G era," *IEEE Communications Magazine,* vol.52, no.2, pp. 90-96, 2014.