

Energy-Efficient Multipath TCP for Quality-Guaranteed Video Over Heterogeneous Wireless Networks

Jiyan Wu , Member, IEEE, Rui Tan , Senior Member, IEEE, and Ming Wang 

Abstract—Prompted by technological advancements in wireless systems and handheld devices, concurrent multipath transfer is a promising solution to stream high-quality mobile videos in heterogeneous access medium. Multipath TCP (MPTCP) is a transport-layer protocol recommended by the Internet Engineering Task Force (IETF) for concurrent data transmission to multi-radio terminals. However, it is still challenging to stream high-quality real-time videos with the existing MPTCP solutions because of the tradeoff between energy efficiency and video quality. To deliver real-time video in an energy-efficient manner, this paper presents a Delay-Energy-quAlity-aware MPTCP (DEAM) solution. First, an analytical framework is developed to characterize the delay-constrained energy-quality tradeoff for multipath video delivery over heterogeneous access networks. Second, a subflow allocation algorithm is proposed to minimize the device energy consumption while achieving target video quality within the imposed deadline. The performance of the proposed DEAM is verified by means of extensive Exata emulations with real-time streaming videos. Evaluation results demonstrate that DEAM achieves appreciable improvements over the reference MPTCP solutions in mobile energy conservation and user-perceived video quality.

Index Terms—Multipath TCP, energy efficiency, real-time video, quality-awareness, heterogeneous wireless networks.

I. INTRODUCTION

THE massive deployments of wireless infrastructures enable individual users to surf the content-rich Internet with various mobile access options, e.g., Wi-Fi, LTE (Long-Term Evolution), HSDPA (High-Speed Download Packet Access), WiMAX (Worldwide Interoperability for Microwave Access), etc. Supported by the latest technological progress, the state-of-the-art mobiles are equipped with different radio interfaces

Manuscript received August 16, 2017; revised December 31, 2017 and July 19, 2018; accepted October 19, 2018. Date of publication November 5, 2018; date of current version May 22, 2019. This research was supported by a research project in Tensor Vision Private Ltd. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Christian Timmerer. (*Corresponding author: Jiyan Wu*)

J. Wu is with the Department of Software Engineering, Tensor Vision Private Ltd., Singapore 139953 (e-mail: wujian@ieee.org).

R. Tan is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: tanrui@ntu.edu.sg).

M. Wang is with the State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: wangming_bupt@bupt.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMM.2018.2879748

to enable simultaneous wireless connections to multiple networks (e.g., the Samsung S5 smartphones [1] and Mushroom bandwidth aggregation products [2]). The prevalence of such multihomed mobile terminals prompts the integration of heterogeneous access medium to provide throughput-demanding mobile applications in future wireless environments [3]–[6].

Mobile video has already taken up the vast majority of Internet data usage due to the proliferation of various multimedia services. As reported in the latest Cisco visual networking index [7], video traffic accounted for 60% of the mobile data transferred over the Internet in 2016 and will exceed 78% by 2021. Since the throughput demand of high-quality streaming applications outpaces the available bandwidth in single wireless networks, it is necessary to consider integrating the heterogeneous access medium for concurrent multipath data transfer, as suggested in existing studies [8]–[15]. The IETF (Internet Engineering Task Force) particularly recommends multipath TCP (MPTCP) [16], [17] as the transport protocol to enable concurrent data transmission over communication networks with multihomed terminals. Distinct from other transport-layer multipath protocols (e.g., SCTP [18] and MPRTP [19]), MPTCP is characterized by TCP-friendliness [20] and firewall penetration capability. Therefore, it is desirable to adopt MPTCP for multihomed video transmission in heterogeneous networks [21]–[23], [70]–[72] (see Fig. 1). In the past few years, MPTCP has been quickly implemented to support various mobile applications, e.g., Siri voice recognition, LTE and Wi-Fi integration [24], [25].

Current battery technologies cannot keep pace with the ever-increasing power demand of energy-intensive multimedia services (e.g., mobile gaming, live sports streaming) on smartphones [26], [27]. In particular, radio communications (e.g., using Wi-Fi/LTE interface) constitute an important part of device energy consumption [67], [68]. Recent studies [28], [29] reveal the vital significance of studying the relationship between energy consumption and user-perceived quality in sustaining multipath video delivery to smartphones. However, *the relationship between energy consumption and video quality in mobile devices* (Proposition 1 in Section III-C) is not considered in the existing MPTCP schemes [30]–[32], [34]. To strike an effective balance, this research models the delay-constrained energy-quality tradeoff for concurrent video transmission with MPTCP to/from multihomed mobile devices. The proposed Delay-Energy-quAlity-aware MPTCP (DEAM) scheme can

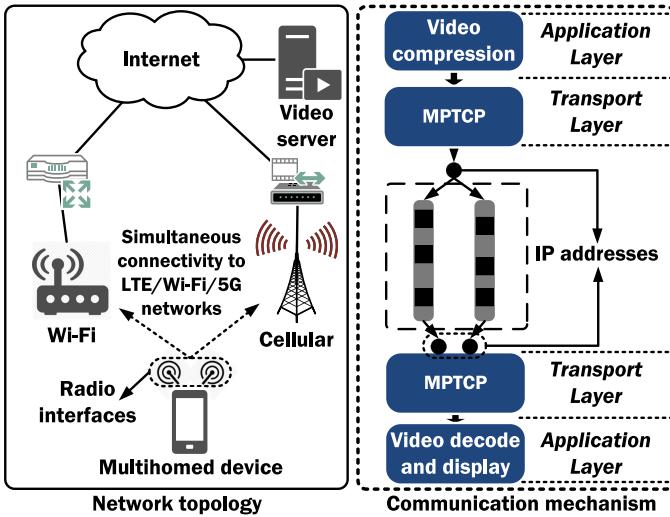


Fig. 1. Multihomed video communication with MPTCP over heterogeneous wireless networks.

effectively exploit the wireless channel diversity and video frame priority to *minimize the mobile energy consumption of mobile devices*.

In particular, the contributions of this research include the following aspects:

- We analyze the tradeoff among delay, energy and quality for concurrent video transmission with MPTCP over multiple wireless access networks.
- We propose a subflow allocation algorithm in MPTCP to minimize energy consumption with regard to the constraints of video quality and transmission deadline.
- We perform extensive semi-physical emulations in the Exata platform involving real-time H.264 video streaming. The evaluation results demonstrate that DEAM achieves performance gains over the reference MPTCP schemes in terms of energy consumption, video PSNR, buffering rate and retransmissions.

The remainder of this paper is organized as follows. Section II reviews related work. In Section III, the theoretical model is developed to analyze the delay-constrained energy-quality tradeoff. Section IV introduces the scheduling algorithms of the proposed DEAM. The performance evaluation is described in Section V. Section VI concludes this work.

II. RELATED WORK

The research efforts related to this study can be broadly divided into three categories: 1) multipath video transmission; 2) energy-efficient Mobile video delivery; 3) performance measurement and optimization on MPTCP.

A. Multipath Video Transmission

In reference [4], an application-layer multipath transmission system is developed for delivering stable live video over heterogeneous wireless networks with fountain code. The authors formulate the optimization problem to maximize video streaming throughput and overcome channel loss. A sub-frame-level video transmission scheme is proposed by Wu *et al.* [9] to split

large video frames for delay performance optimization of HD (high-definition) video streaming over multihomed mobile devices. A probabilistic multipath transmission (PMT) scheme was proposed in reference [35] to deliver video traffic over multiple available paths based on the probability generation function of end-to-end packet delay. In [36], the authors developed a dynamic rate-allocation-based Joint Source-Channel Coding (JSCC) to optimize the mobile streaming quality over multihomed mobiles. In reference [3], the video distortion model was introduced into SCTP to optimize mobile video streaming quality over heterogeneous wireless networks. A real-time adaptive algorithm was developed by Xing *et al.* [5] for video streaming over multiple access networks based on Markov decision process. A loss-resilient bandwidth aggregation scheme was proposed in [11] to spread FEC packets over multiple wireless networks to mitigate burst losses.

The motivation of the above studies is to integrate heterogeneous network resources for improving the quality of video streaming. This paper addresses the problem of multipath video transmission by taking advantage of MPTCP.

B. Energy-Efficient Mobile Video Delivery

Reference [27] reviews recent studies on energy-efficient video delivery. A download scheduling algorithm based on crowd-sourced video viewing statistics was proposed by Hoque *et al.* [26] to reduce aggressive prefetching and tail energy consumption. In reference [37], the authors proposed an online algorithm for energy conservation in user abandonment scenarios, and an offline solution to estimate minimal power consumption. Bui *et al.* [38] designed and implemented the GreenBag solution, which includes the load balancing, segment management and energy-aware mode control for delivering real-time video to multi-radio devices. Ismail *et al.* [28] develop an energy management sub-system for mobile terminals to sustain multipath video streaming. In reference [29], an energy- and content-aware packet scheduling approach was proposed for video upload service on multihomed devices. Huang *et al.* [42] presented a close examination of power consumption of radio interfaces on smartphones. Aparicio-Pardo *et al.* [69] proposed a green video control plane with fixed-mobile convergence and cloud-RAN (Radio Access Networks) to optimize energy consumption and QoE (quality-of-experience).

Energy consumption of mobile video streaming is a critical problem as investigated in the existing studies. However, the energy efficiency of video communication with MPTCP still remains to be explored.

C. Performance Optimization on MPTCP

The energy-efficient MPTCP scheme proposed in reference [30] models the throughput-energy tradeoff for path selection and congestion control. However, the authors in [30] did not consider video quality, which is distinct from network-level throughput [3]. The fountain code-based multipath TCP (FMTCP) proposed by Cui *et al.* [31] leverages rateless fountain coding to combat packet loss and path heterogeneity in multipath data transfer. Li *et al.* [32] proposed the SC-MPTCP scheme, which introduces the linear systematic coding to

TABLE I
DIFFERENCES BETWEEN THE PROPOSED DEAM AND THE EXISTING WORKS [3], [9], [11], [34], [36], [38]

| Solution | protocol layer | data protection | data/rate allocation | energy-awareness |
|---------------|----------------------------|---------------------------------------|-----------------------------------|------------------|
| DEAM | transport layer (MPTCP) | delay and energy-aware retransmission | quality-aware energy minimization | ✓ |
| CMT-DA [3] | transport layer (SCTP) | delay and loss-aware retransmission | distortion minimization | ✗ |
| SFL [9] | application layer | ✗ | water-filling | ✗ |
| LTBA [11] | application layer | FEC (Reed-Solomon codes) | bandwidth-based | ✗ |
| FRA-JSCC [36] | application-physical layer | FEC (Reed-Solomon codes) | greedy search | ✗ |
| GreenBag [38] | application layer | retransmission | goodput-based | ✓ |
| ADMIT [34] | transport layer (MPTCP) | FEC & retransmission | utility maximization | ✗ |

address the problem of tolerating path heterogeneity under receiver buffer limitation. The FEC coding redundancy is provisioned into both proactive and reactive data. A cross-layer design named MPTCP-MA was developed in reference [33] and uses MAC-layer information to locally estimate path quality and connectivity. Chen *et al.* [21] performed a measurement study of MPTCP performance over cellular and Wi-Fi networks to investigate the impact of path diversity on application-level metrics. Han *et al.* [39] proposed algorithms to split flows over multiple paths in the Internet by using minimal congestion feedback signals. Wu *et al.* [34] proposed a quality-driven MPTCP scheme called ADMIT that leverages FEC (Forward Error Correction) coding to stream high-quality real-time video. However, the energy consumption of different radio interfaces is not considered in ADMIT [34] and this significantly degrades the energy efficiency of multipath video streaming. A cross-layer scheduler was introduced in reference [70] to address the abrupt throughput drops and optimize video streaming quality. Han *et al.* [71] develop the MP-DASH to strategically schedule video chunk's delivery and satisfy user preference of network interface. Chen *et al.* [73] proposed an energy-aware MPTCP (eMPTCP) that considers the energy consumption of radio interfaces to improve data throughput and energy efficiency. Distinct from the above studies [30]–[34] on MPTCP, this paper models and leverages the delay-constrained tradeoff between energy consumption and received video quality.

The main differences between the proposed DEAM and the existing works [3], [9], [11], [34], [36], [38] are summarized in Table I. In our previous work [66], we investigate the energy-quality tradeoff for multihomed video communication with MPTCP. This paper provides extensions in the system model, scheduling algorithm and performance evaluation parts. To the best of our knowledge, the presented DEAM is the first MPTCP solution that exploits the delay-constrained energy-quality tradeoff to optimize video delivery in heterogeneous wireless access networks.

III. SYSTEM MODEL

A. System Overview

Fig. 2 presents the system overview of the proposed MPTCP solution. This research considers the problem of end-to-end wireless video communication with MPTCP over multiple available access networks. The objective of the proposed DEAM scheme is to minimize the device energy consumption under quality and delay constraints imposed by real-time video

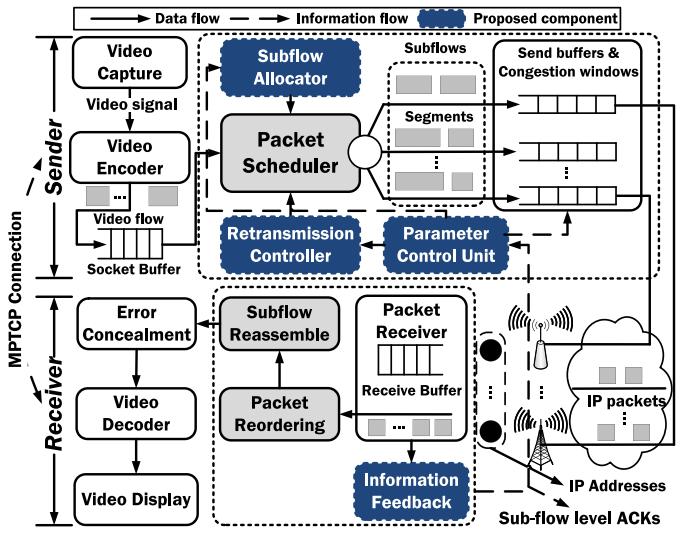


Fig. 2. System overview of the proposed DEAM (Delay-Energy-quAlity aware MPTCP) scheme.

applications. The video packets are distributed onto multiple subflows over the available communication paths. Each subflow consists of several transmission segments. The decision blocks in the system framework are implemented at both communication terminals. The key contribution of this research is the subflow allocation module of MPTCP. The components at the sender side include the parameter control unit, subflow allocator, and packet retransmission controller.

The information feedback unit at the receiver side is used for periodically reporting the channel status and subflow-level acknowledgments (ACKs) to the sender side. The packets may be received by the destination out of order because of the path heterogeneity in different access networks. These out-of-order packets will be reordered to reconstruct the original video traffic. To alleviate the playback disruptions caused by packet drops, the error concealment method using frame copy is implemented at the receiver side.

The key research issue in energy and quality optimization of MPTCP is the subflow allocation (i.e., video packet distribution over the available communication paths). This optimization problem involves the analytical models of wireless access network [3], video distortion [40], and energy consumption [41], [42]. To present detailed descriptions of the system framework, we briefly introduce these mathematical models in this section. The basic mathematical symbols used throughout this work are explained in Table II.

TABLE II
DEFINITIONS OF MATHEMATICAL SYMBOLS

| Symbol | Definition |
|--------------------------|---|
| \mathbb{P}, \mathbb{E} | the probability value, expectation value. |
| \mathcal{P}, p | the set of communication paths, a path element. |
| T, P | the delay constraint, number of available path. |
| \mathbf{R}, R_p | the subflow allocation vector, an element. |
| RTT_p | the round-trip time of p . |
| μ_p, ν_p | the available, residual bandwidth of p . |
| D, \bar{D} | the end-to-end distortion, quality requirement. |
| π_p^B | the packet loss rate of p . |
| Π_p | the effective loss rate over p . |
| π_p^t, π_p^o | the transmission, overdue loss rate over p . |
| G/B | the Good/Bad state of p . |
| ξ_p^G | the state transition probability of p from B to G . |
| $Q_p^{(i,j)}(t)$ | the transition probability from i to j in time t . |
| n, n_p | the total number of packets, dispatched onto p . |
| D_p | the end-to-end delay over path p . |
| \mathcal{U}, U_p | the system utility matrix, an element. |
| e_p | the energy consumption parameter of p . |

B. Model Description

Wireless Access Network Model

For this mathematical framework, we consider a heterogeneous wireless network integrating P access networks between communication terminals. Each end-to-end link indicates an independent path in MPTCP and can be established by binding a pair of IP addresses. A communication path $p \in \mathcal{P}$ is considered to be an independent connection uncorrelated with any others. The physical properties of each path include the available bandwidth μ_p , round trip time RTT_p , and packet loss rate π_p^B . Reference [3] presents the detailed descriptions of these network parameters.

The packet loss pattern is analyzed with the Gilbert loss model [43], which can be expressed as a stationary continuous-time Markov chain with two states. On time t , the state $\mathcal{X}_p(t)$ value is selected from two options: G (Good) or B (Bad). There is no packet loss in the Good path status [44], [45]. If a packet is delivered at time t and $\mathcal{X}_p(t) = G$, the packet can be successfully delivered to the destination. Otherwise, the packet is lost if $\mathcal{X}_p(t) = B$. Let π_p^G and π_p^B denote the stationary probabilities that path p is good or bad. Assume ξ_p^G and ξ_p^B denote the transition probability from B to G and G to B , respectively. This research employs two system-dependent parameters to characterize the packet loss behavior: 1) the packet loss rate π_p^B and 2) the average loss burst length $1/\xi_p^B$. We can have $\pi_p^G = \xi_p^G / (\xi_p^B + \xi_p^G)$ and $\pi_p^B = \xi_p^B / (\xi_p^B + \xi_p^G)$.

Effective Loss Rate

The effective loss rate denotes the combined loss ratio of dropped and overdue packets over all communication paths.

Specifically, the following equation is presented to estimate the effective loss rate over p

$$\Pi_p = \pi_p^t + (1 - \pi_p^t) \cdot \pi_p^o, \quad (1)$$

where π_p^t denotes the transmission loss rate, and π_p^o represents the overdue loss rate. The definitions for these terminologies are listed as follows.

Definition 1: The transmission loss rate denotes the percentage of lost packets in the end-to-end data transfer. Such packet losses can be caused by the channel fading, link congestion, external interference, etc.

Definition 2: The overdue loss rate indicates the probability of packets received by the destination after the deadline (T) imposed by the video application.

First, we present the mathematical analysis for the transmission loss rate π_p^t based on the Gilbert loss model and continuous-time Markov chain.

In the scheduling process of MPTCP, the input data traffic (e.g., the video frames in a GoP with total size S) is divided into multiple subflows ($S_p = (R_p \cdot S)/R, p \in \mathcal{P}$) and each subflow is dispatched onto a different path. The segments in each subflow are split into IP (Internet protocol) packets at the network layer. Therefore, the number of packets on each path can be estimated with $n_p = \lceil S_p / MTU \rceil$, where $\lceil x \rceil$ represents the smallest integer larger than x and MTU denotes the maximum transmission unit size. In this analytical framework, we assume the packets over each path are evenly spread with interval ω_p to approach the transmission loss rate. Suppose the notation c_p represents a n_p -tuple, which represents a specific failure configuration of path p . $c_p^i = B, p \in \mathcal{P}, 1 \leq i \leq n_p$ indicates the i th packet assigned to path p is lost and vice versa. We can obtain the equation for the transmission loss rate π_p^t by considering all possible configurations of c_p :

$$\pi_p^t = \frac{1}{n_p} \cdot \sum_{\text{all } c_p} L(c_p) \cdot \mathbb{P}(c_p). \quad (2)$$

The notation “all c_p ” represents all the possible combinations of c_p ; $L(c_p)$ denotes the number of lost packets on path p . Equation (2) is used to estimate the *expected value* of the transmission loss rate over path p , i.e., $\pi_p^t = \frac{\text{expected number of lost packets}}{\text{total number of packets}}$. The denominator of Equation (2) is $\sum_{\text{all } c_p} L(c_p) \cdot \mathbb{P}(c_p)$, and the numerator is n_p (i.e., the total number of packets scheduled onto path p). As $\sum_{\text{all } c_p} L(c_p) \cdot \mathbb{P}(c_p)$ denotes the expected number of lost packets (caused by transmission losses) over path p , $\frac{\sum_{\text{all } c_p} L(c_p) \cdot \mathbb{P}(c_p)}{n_p}$ represents the expected transmission loss rate. $L(c_p)$ can be expressed as

$$L(c_p) = \sum_{i=1}^{n_p} \mathbb{1}_{\{c_p^i=B\}}.$$

Suppose $\mathbb{P}(c_p^i)$ denotes the failure probability of path failure on p when transmitting the i th packet. The state transition probability of path p from state i to j in time ω_p is expressed with the symbol $Q_p^{(i,j)}(\omega_p)$, i.e.,

$$Q_p^{(i,j)}(\omega_p) = \mathbb{P}[\mathcal{X}_p(\omega_p) = j | \mathcal{X}_p(0) = i].$$

Based on the transient property of the continuous-time Markov chain, the following equation presents the state transition matrix:

$$\begin{aligned}\mathbf{Q}_p^{(G,G)}(\omega_p) &= \pi_p^G + \pi_p^B \cdot \exp [-(\xi_p^B + \xi_p^G) \cdot \omega_p], \\ \mathbf{Q}_p^{(G,B)}(\omega_p) &= \pi_p^B - \pi_p^B \cdot \exp [-(\xi_p^B + \xi_p^G) \cdot \omega_p], \\ \mathbf{Q}_p^{(B,G)}(\omega_p) &= \pi_p^G - \pi_p^G \cdot \exp [-(\xi_p^B + \xi_p^G) \cdot \omega_p], \\ \mathbf{Q}_p^{(B,B)}(\omega_p) &= \pi_p^B + \pi_p^G \cdot \exp [-(\xi_p^B + \xi_p^G) \cdot \omega_p].\end{aligned}$$

Therefore, the value of $\mathbb{P}(c_p)$ can be estimated with the following equation.

$$\mathbb{P}(c_p) = \prod_{i=1}^{n_p} \mathbb{P}(c_p^i) = \pi_p^{c_p^i} \cdot \prod_{i=1}^{n_p-1} \left[\mathbf{Q}_p^{(c_p^i, c_p^{i+1})}(\omega_p) \right].$$

Finally, π_p^t can be obtained with

$$\pi_p^t = \left\lceil \frac{MTU}{S_p} \right\rceil \cdot \sum_{\text{all } c_p} \pi_p^{c_p^i} \cdot \prod_{i=1}^{\left\lceil \frac{S_p}{MTU} \right\rceil - 1} \left[\mathbf{Q}_p^{(c_p^i, c_p^{i+1})}(\omega_p) \right]. \quad (3)$$

The overdue loss rate (π_p^o) is dependent on the end-to-end packet transmission delay (D_p). This latency is dominated by the queueing delay at the bottleneck link of communication path p , and asymptotically follows the exponential distribution [6], [46]. Therefore, the overdue loss rate π_p^o is expressed as follows:

$$\pi_p^o = \exp \left[-\frac{T}{\mathbb{E}(D_p)} \right], \quad (4)$$

where $\mathbb{E}(\cdot)$ represents the expectation value. Conventionally, the value of $\mathbb{E}(D_p)$ is estimated based on a large collection of end-to-end delay statistics. To develop a solution for online estimation, this research employs a model to approximate the average packet delay. Suppose the residual bandwidth of p is denoted with symbol ν_p . We can have $\nu_p = \mu_p - R_p$.

Since the assigned subflow rate on each path is often close to the available bandwidth, the mean end-to-end delay generally increases because of network congestion. A fractional function is employed to approximate the end-to-end delay of the allocated subflow $S_p = \frac{R_p \cdot S}{R}$ over path p , i.e.,

$$\mathbb{E}(D_p) = \frac{S_p}{\mu_p} + \frac{\rho_p}{\nu_p},$$

where ρ_p represents the available source of communication path p . The value of ρ_p can be estimated from the latest measured values of the path status information [6], [47]

$$\rho_p = \frac{\nu'_p \cdot RTT_p}{2}.$$

If ν'_p is equal to the latest observed residual bandwidth of path p , i.e., $\nu'_p = \nu_p$, the one-way delay is $RTT_p/2$. The mathematical formula for the overdue loss rate is presented as follows.

$$\pi_p^o = \exp \left(-\frac{2 \times T \cdot \nu_p \cdot \mu_p}{\nu'_p \cdot RTT_p \cdot \mu_p + 2 \times \nu_p \cdot R_p} \right). \quad (5)$$

The effective loss rate can be obtained with Equations (3), (4) and (5).

Video Distortion Model

In this analytical framework, a widely used end-to-end video distortion model [40] is adopted to optimize the real-time streaming quality. The objective video quality in real-time applications is dependent on the end-to-end distortion (D). In particular, D is expressed as the sum source distortion (D_{src}) and channel distortion (D_{chl}), i.e.,

$$D = D_{src} + D_{chl}. \quad (6)$$

This mathematical model indicates that the quality of real-time video streaming is determined by 1) the distortion caused by the data compression of the media contents and 2) the distortion due to the packet drops in end-to-end video transport. Specifically, the source distortion mainly depends on the encoding bitrate (R) and video sequence content. These parameters largely impact the coding efficiency of the video codec (e.g., a lower encoding bitrate leads to higher source distortion for the same video sequence). The channel distortion is mainly caused by the effective loss rate (Π), which is defined as the overall loss rate of data packets that encounter channel errors/losses or overdue arrivals [3].

D_{chl} is generally proportional to the number of dropped video frames. Therefore, the end-to-end video distortion can be mathematically expressed [in units of mean square error (MSE)] as follows [47]:

$$D = \frac{\alpha}{R - R_0} + \beta \cdot \Pi, \quad (7)$$

where α , R_0 and β are parameters depending on the specific video codec and sequences. According to Table I in reference [40], α is defined as the RD (rate distortion) factor, and its value is dependent on the video encoder and intra-rate (percentage of macroblocks). β represents the sensitivity of the video sequence to packet losses [47]. These parameters can be estimated from three or more trial encodings using nonlinear regression techniques [47]. To enable fast subflow allocation of abrupt changes in the video content, these parameters can be updated for each GoP in the encoded video sequence (typically every 0.5 seconds) [47]. It is feasible to perform online estimation of the distortion parameters since powerful servers are currently prevalent.

Energy Consumption Model

The energy models proposed in references [41], [42] are employed to analyze the energy/power consumption of mobile devices. These models mainly consider the ramp, transfer and tail energy. The mobile energy and power consumption are measured [in units of Joules (J)] based on the factors of packet length, signaling frequency and data transfers. The data transfer energy is considered in our analytical framework as it constitutes the majority of energy dissipation for multimedia transmission to mobiles [26]. The energy consumption parameter e_p for a specific communication path p is defined as the consumed energy for delivering the same dat volume (J/Kbps). The total energy consumption for the given subflow allocation vector $\mathbf{R} = \{R_p\}_{p \in P}$ of all communication paths can then be

expressed as

$$\mathbf{E} = \sum_{p \in P} R_p \cdot \mathbf{e}_p. \quad (8)$$

Extensive measurement studies [26], [42] reveal that the energy consumed for transmitting the same amount of data through a Wi-Fi network is lower than that with 3G (e.g., WCDMA) and LTE.

C. Analysis of Energy-Quality Tradeoff

The end-to-end distortion with regard to the subflow allocation vector can be estimated with the following equation

$$D = \frac{\alpha}{R - R_0} + \beta \cdot \frac{\sum_{p \in P} R_p \cdot \Pi_p}{\sum_{p \in P} R_p}, \quad (9)$$

Therefore, the end-to-end distortion D is proportional to the effective loss rate Π_p . This loss rate indicates the quality of the communication path, and thus is different from the round trip time, available bandwidth, or packet loss rate. In particular, the following proposition is presented to explain the delay-constrained energy-quality tradeoff for multipath video transmission over heterogeneous wireless networks.

Proposition 1: For multipath video transmission to mobile devices over different wireless access networks (e.g., Wi-Fi, 3G, LTE), the received video quality is generally proportional to the device energy consumption, i.e., higher video quality (lower video distortion) requires more energy consumption for video data transfer using radio interfaces.

Proof: As shown in existing studies [3], [4], [48], cellular networks exhibit higher resilience in sustaining user mobility than Wi-Fi networks. Therefore, the effective loss rate of data through a Wi-Fi network is higher than that using cellular networks, i.e., $\Pi_W > \Pi_C$. In addition, the mobile energy to transport the same data volume using the Wi-Fi radio interface is lower than that through a cellular network. Assume there are two subflow allocation schemes with assignment vectors of $a = [R_W^a, R_C^a]$ and $b = [R_W^b, R_C^b]$ for the same video flowrate R ($R_W^a < R_W^b, R_C^a > R_C^b$). The mobile energy consumption is compared as follows.

$$R_W^a \cdot \mathbf{e}_W + R_C^a \cdot \mathbf{e}_C > R_W^b \cdot \mathbf{e}_W + R_C^b \cdot \mathbf{e}_C \Rightarrow \mathbf{E}^a > \mathbf{E}^b.$$

Furthermore, the end-to-end distortion values are compared as follows:

$$\begin{aligned} \frac{R_W^a \cdot \Pi_W + R_C^a \cdot \Pi_C}{R} &< \frac{R_W^b \cdot \Pi_W + R_C^b \cdot \Pi_C}{R} \\ \Rightarrow \Pi^a &< \Pi^b \Rightarrow D^a < D^b. \end{aligned}$$

For another case of subflow allocation, we assume there are two subflow allocation schemes $a = [R_W^a, R_C^a]$ and $b = [R_W^b, R_C^b]$ for the same video flowrate R ($R_W^a > R_W^b, R_C^a < R_C^b$). The comparison of energy consumption is as follows.

$$R_W^a \cdot \mathbf{e}_W + R_C^a \cdot \mathbf{e}_C > R_W^b \cdot \mathbf{e}_W + R_C^b \cdot \mathbf{e}_C \Rightarrow \mathbf{E}^a < \mathbf{E}^b.$$

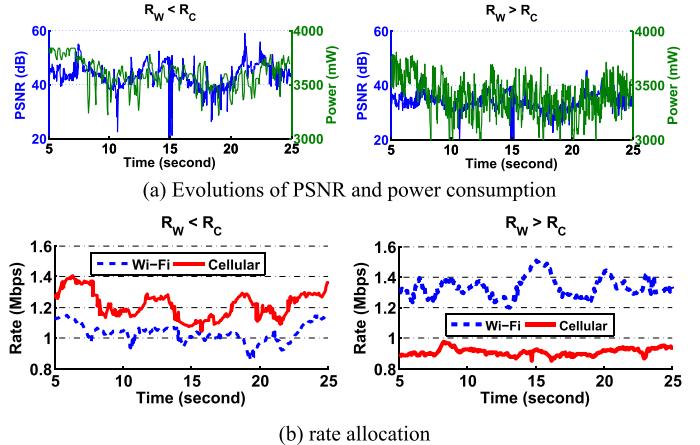


Fig. 3. Video flow allocation and power consumption over both Wi-Fi and cellular networks.

For the same rate allocation vectors, the end-to-end distortion values D are compared as follows

$$\begin{aligned} \frac{R_W^a \cdot \Pi_W + R_C^a \cdot \Pi_C}{R} &< \frac{R_W^b \cdot \Pi_W + R_C^b \cdot \Pi_C}{R} \\ \Rightarrow \Pi^a &> \Pi^b \Rightarrow D^a > D^b. \end{aligned}$$

Based on the above derivations, we can conclude that delivering more video frames over the cellular interface achieves better video quality in terms of PSNR but also incurs higher device energy consumption.

The following example is presented to explain the delay-constrained energy-quality tradeoff in detail.

Example 1: Suppose a video flow is concurrently transmitted through both Wi-Fi and cellular to the multihomed mobile terminal. The selected video test sequence is the river bed (medium motion level and scene complexity) and it is encoded at 2.5 Mbps, 30 fps. The GoP structure is IPP...P to ensure low-delay compression. Fig. 3(a) shows the evolution of the power consumption and PSNR per video frame during the interval of [5, 25] seconds for both cases of $R_W > R_C$ and $R_W < R_C$. We can observe that the variations in the PSNR values for the received video frames track closely to the power consumption as shown in Fig. 3(b). These results verify the conclusion in Proposition 1.

IV. PROBLEM FORMULATION AND SCHEDULING ALGORITHMS

A. Problem Formulation

The objective optimization problem can be stated as follows. Given the feedback channel status $\{\text{RTT}_p, \mu_p, \pi_p^B\}_{p \in P}$, quality requirement \bar{D} , delay constraint T , and input video frames, find the optimal subflow allocation vector $\mathbf{R} = \{R_p\}_{p \in P}$ to minimize the mobile energy consumption $\mathbf{E} = \sum_{p \in P} R_p \cdot \mathbf{e}_p$. Mathematically, this *energy minimization* problem with quality

and delay constraints can be formulated as follows.

$$\{R_p\}_{p \in P} = \arg \min \left\{ \sum_{p \in P} R_p \cdot \mathbf{e}_p \right\}, \quad (10)$$

$$\sum_{p \in P} R_p \cdot \Pi_p = \frac{R}{\beta} \cdot \left(\bar{D} - \frac{\alpha}{R - R_0} \right), \quad (11a)$$

$$\text{subject to } R_p \leq \mu_p \cdot (1 - \pi_p^B), \quad (11b)$$

$$\frac{S_p}{\mu_p} + \frac{\nu'_p \cdot RTT_p}{2 \times \nu_p} \leq T. \quad (11c)$$

where $\Pi_p = \pi_p^t + (1 - \pi_p^t) \cdot \pi_p^o$,

$$\pi_p^t = \left\lceil \frac{MTU}{S_p} \right\rceil \cdot \sum_{\text{all } c_p} \pi_p^{c_p^i} \cdot \prod_{i=1}^{\left\lceil \frac{S_p}{MTU} \right\rceil - 1} \left[Q_p^{(c_p^i, c_p^{i+1})}(\omega_p) \right],$$

$$\pi_p^o = \exp \left(-\frac{2 \times T \cdot \nu_p \cdot \mu_p}{\nu'_p \cdot RTT_p \cdot \mu_p + 2 \times \nu_p \cdot R_p} \right),$$

Condition 11a states the target video quality (\bar{D}) imposed by the real-time applications (e.g., average, good, excellent level). The bandwidth and delay constraints for each communication path are regulated in conditions 11b and 11c, respectively. The objective function (10) is an optimization problem with both equality and inequality constraints. We prove the NP hardness of problem (10) by reducing the partition problem [49], and the proof is presented in the supplementary file. Consequently, there is no optimal solution with polynomial time complexity. This section presents a low-complexity heuristic algorithm for online subflow allocation.

B. Subflow Allocation

The solution procedure of the proposed subflow allocation scheme includes the following steps: 1) the adaptation of the video traffic rate to guarantee the imposed quality constraint D and 2) subflow allocation to minimize energy consumption of multipath data transfer to mobile devices.

Quality-Guaranteed Load Adaptation : The video traffic load adaptation procedure is based on Proposition 1 in Section III-C, i.e., higher video quality requires more mobile energy consumption for data transfer. Therefore, this section presents an algorithm to appropriately reduce the video traffic rate based on the quality requirement D and video frame priority. Specifically, the video frames are identified by different types [e.g., predicted (P), intra (I), bidirectional (B) frames], inter-frame dependency, etc. The payload length and frame type information can be analyzed by reading the header information of the NALU (network abstract layer unit) [50]. Since the proposed framework is a completely transport-layer solution and the video encoding parameters (e.g., bit rate, frame rate, quantization parameter) are not within the control scope, we use the selective video frame drop to reduce the transmission rate with regard to the video quality requirement. In the existing studies (e.g., [72]),

Algorithm 1: Quality-Guaranteed Load Adaptation

```

Input: Delay constraint  $T$ , network parameters
        $\{RTT_p, \mu_p, \pi_p\}_{p \in P}, \bar{D}, R, F;$ 
Output: Updated video transmission rate  $R$ ;
1  $D = \frac{\alpha}{R - R_0} + \beta \cdot \frac{\sum_{p \in P} R_p \cdot \Pi_p}{\sum_{p \in P} R_p};$ 
2 while  $D \leq \bar{D}$  do
3   Select the video frame  $f$  with the lowest weight  $\omega_f$  from
      the set of all the frames  $F$ ;
4   Update the total transmission rate  $R$  by considering the
      drop of the frame  $f$ ;
5    $\pi_p^t = \left\lceil \frac{MTU}{S_p} \right\rceil \cdot \sum_{\text{all } c_p} \pi_p^{c_p^i} \cdot \prod_{i=1}^{\left\lceil \frac{S_p}{MTU} \right\rceil - 1} \left[ Q_p^{(c_p^i, c_p^{i+1})}(\omega_p) \right];$ 
6    $\pi_p^o = \exp \left( -\frac{2 \times T \cdot \nu_p \cdot \mu_p}{\nu'_p \cdot RTT_p \cdot \mu_p + 2 \times \nu_p \cdot R_p} \right);$ 
7    $\Pi_p = \pi_p^t(R_p) + [1 - \pi_p^t(R_p)] \cdot \pi_p^o(R_p);$ 
8    $R_p = \frac{\mu_p \cdot (1 - \pi_p^B) \cdot R}{\sum_{p \in P} \mu_p \cdot (1 - \pi_p^B)}, \forall p \in P;$ 
9    $D = \frac{\alpha}{R - R_0} + \beta \cdot \frac{\sum_{p \in P} R_p \cdot \Pi_p}{\sum_{p \in P} R_p};$ 
10 end

```

frame dropping is proved to be an effective solution to adapting traffic rate.

Dropping higher-priority video frames (e.g., I frames) can reduce more traffic load than discarding the lower-weight frames. However, this operation will result in the decoding failure of subsequent P frames depending on the I frame. Algorithm 1 obtains the minimal traffic rate R to satisfy the upper bound of video distortion \bar{D} , by dropping the video frames with lower priority. The authors in [13] proved that the loss-free bandwidth ($\mu_p \cdot (1 - \pi_p^B)$) can indicate the path quality. In the proposed algorithm, we set the initial rate assignment to be proportional to this parameter. We assume there are a total of F frames to be scheduled and each frame $f \in F$ is associated with a weight w_f . The per-frame weight in this paper is defined according to reference [28], [29], [51], i.e., $w_f = 5$ for I frames, 3 for P frames and 2 for B frames. In the emulations of this paper, we use the GoP structure of IP...PP, i.e., without B frames. For the same type of frames (e.g., P frames in a GoP), we first drop the frame in the tail of the GoP. Algorithm 1 presents the video traffic rate adaptation process with regard to the video distortion constraint. A critical problem in dropping video frames to balance energy-quality tradeoff is to guarantee the video quality requirement. As shown in the main while loop (Line 2–10) of Algorithm 1, we limit the estimated video distortion to be smaller than the quality requirement \bar{D} to achieve this objective. If the estimated distortion is greater than the quality requirement, the while loop terminates and no more video frames will be dropped.

This load adaptation model can be implemented in the packet scheduling module of MPTCP, and polynomial time complexity is feasible for online operation. The video frame type can be identified by reading the payload header according to the specifications in RFC 3984 [50].

The worst-case time complexity of Algorithm 1 is $O(F \cdot (F + n \cdot k))$, where F denotes the GoP size (i.e., the number of frames in a GoP). There are no more than F times of operations in the **while** loop of Algorithm 1 to drop the video frames. The

time complexity to estimate the transmission loss rate π_p^t is $n \cdot k$. Thus, the time complexity of Algorithm 1 is $O(F \cdot (F + n \cdot k))$.

Subflow Allocation Algorithm : Since the energy consumption E is dependent on the subflow rate allocation, piecewise linear approximation can be obtained with an univariate function. This approximation can be achieved by dividing the interest region of each univariate function into an adequate number of non-overlapped small intervals. It is then possible to adopt convex piecewise linear function to approach the objective function on each hypercube. This objective function can be presented as an arbitrary univariate function $l(\cdot)$. The feasible solutions for energy minimization are in the possible flowrate allocation vectors. Therefore, the critical issue is to select the appropriate breakpoints and determine whether they are inflection points for the piecewise linear approximation. The multiple knapsack problem (10) is addressed with piecewise linear approximation and utility maximization theory. The motivation to adopt this mathematical framework is summarized as follows:

Proposition 2: The subflow allocation scheme based on utility maximization theory and piecewise linear approximation can approximate the minimal energy consumption under quality and delay constraints.

Proof: The proof is provided in the supplementary file.

The proposed subflow allocation algorithm may assign excessive packets to the communication paths with higher quality in order to achieve the target video quality (\bar{D}). The quality level of a communication path is proportional to its transport capability. However, this method may lead to link congestion and load imbalance during multipath data transfer. Similarly, the energy-minimized allocation may also incur overload problems since most of the video packets are dispatched to the path with lowest mobile power dissipation. To mitigate serious load imbalance issues, a load imbalance parameter L_p is introduced to represent the load level of path p and it can be calculated with the formula of $L_p = \frac{\mu_p \cdot (1 - \pi_p) - R_p}{\sum_{i=1}^P [\mu_i \cdot (1 - \pi_i) - R_i] / P}$, where $\mu_p \cdot (1 - \pi_p)$ represents the ‘loss-free’ bandwidth of path p . Path p is overloaded if the value of L_p is obviously higher than a threshold limit value (TLV) [3], [47]. The initial subflow allocation for each path is set to be proportional to the available bandwidth, i.e., $R_p = (R \cdot \mu_p) / \sum_{i=1}^P \mu_i$. Let ΔR_p denote the rate variation over path p at each iteration and $R_p + \Delta R_p$ represent the transition of the next allocation. The utility of this transition can be expressed as $U_p(R_p) = \frac{\varphi(R_p + \Delta R_p) - \varphi(R_p)}{\Delta R_p \cdot e_p}$, where $\varphi(\cdot)$ represents the approximate linear function for D_{total} in the interval $[R_p, R_p + \Delta R_p]$. The utility matrix can be expressed as $\mathcal{U} = \{U_p\}_{p \in \mathcal{P}}$. In each cycle, the developed flowrate allocation scheme calculates the R_p value that entails the highest system utility value, i.e., $\mathbf{R} = \text{argmax}_{\{R_p\}_{p \in \mathcal{P}}} \{\sum_{p \in \mathcal{P}} U_p(R_p)\}$.

After obtaining a feasible solution as the rate allocation vector to minimize the mobile energy consumption, an inter-path allocation scheme is developed to improve the system utility. For a transmission path p , this algorithm seeks another communication path p' to transfer part of the allocated rate if the aggregated system utility can be improved. If the resources of path p are exhausted, a different path will be selected to release some of the available channel resources. This operation continues until the system utility cannot be optimized or the available network

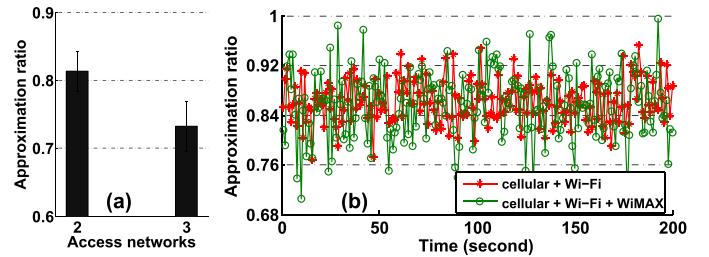


Fig. 4. Approximation ratio of the proposed algorithm: (a) average values, (b) instantaneous values.

resources are depleted. The outline of the subflow allocation algorithm based on utility maximization and linear approximation is presented in Algorithm 2. Specifically, we present the line-by-line algorithm explanations as follows

- Lines 1–8: The foreach loop determines the rate allocation R_p for each communication path with the approximate linear function $\varphi(\cdot)$.
- Line 2: Initialize the rate allocation R_p according to the available bandwidth, and calculate the utility value U_p of path p .
- Line 3: Estimate the load imbalance factor of path p with Equation (12).
- Lines 4–7: The while loop adjusts the rate allocation (with the step of ΔR) of communication path p under the constraints of the load parameter and available bandwidth.
- Line 8: Call the inter-path allocation to improve the obtained rate allocation solution.
- Line 10: The rate allocation vector to achieve the highest system utility can be obtained after the foreach loop.
- Lines 11–20: The inter-path allocation procedure improves the feasible solution.
- Line 12: State the condition for the inter-path allocation, i.e., the load balance and bandwidth constraints.
- Line 13: Update the free network resources of communication path p .
- Line 14: Find another path p' that also satisfies the load balance and bandwidth constraints.
- Line 15: Calculate and compare the system utility after the inter-path rate allocation.
- Lines 16–17: Update the rate allocation and path resources if the system utility can be improved.

The time complexity of Algorithm 2 is $O(P + P^2 \cdot \frac{R}{\Delta R})$. It takes P iterations to determine the rate allocation R_p for each communication path. The time complexity of the improvement steps (Lines 11–18) is $O(P^2 \cdot \frac{R}{\Delta R})$. Therefore, the time complexity of Algorithm 2 is $O(P + P^2 \cdot \frac{R}{\Delta R})$.

The approximation ratio values of the proposed subflow allocation algorithm are presented in Fig. 4. Specifically, we compare the performance of the proposed algorithm with the optimal solution. The optimal solution is achieved with an omniscient scheduler that knows the entire future of the process, and such an omniscient scheduler is conventionally called an **oracle**. The evaluation results indicate that the approximation

Algorithm 2: Subflow Allocation

Input: $\{RTT_p, \mu_p, \pi_p\}_{p \in \mathcal{P}}$, T, $TLV = 1.2$, $\Delta R = 0.05 \times R$;
Output: $R = \{R_p\}_{p \in \mathcal{P}}$;

foreach element p in available path set \mathcal{P} **do**

- 2 $U_p(R_p) = \frac{\varphi(R_p + \Delta R_p) - \varphi(R_p)}{\Delta R_p \cdot e_p}$, where $R_p = \frac{\mu_p \cdot R}{\sum_{i=1}^P \mu_i}$;
- 3 $L_p = \frac{\mu_p \cdot (1 - \pi_p) - R_p}{\sum_{i=1}^P [\mu_i \cdot (1 - \pi_i) - R_i] / P}$;
- 4 **while** $(R_p \leq \mu_p \cdot (1 - \pi_p^B)) \& \& (L_p \leq TLV)$ **do**
- 5 $\Delta R_p = \Delta R_p / U_p$, $R_p \leftarrow R_p + \Delta R_p$;
- 6 Update the approximate function $\varphi(R_p)$;
- 7 **end**
- 8 Call the inter-path allocation procedure (Line 12- 19) to improve the obtained rate allocation vector;
- 9 **end**
- 10 $R = \underset{R_p}{\operatorname{argmax}} \left\{ \sum_{p=1}^P U_p(R_p) \right\}$;

11 Inter-path allocation for solution improvement:

- 12 **if** $(R_p \leq \mu_p \cdot (1 - \pi_p^B)) \& \& (L_p \leq TLV)$ **then**
- 13 Update free resources of path p ;
- 14 find other path p' subject to:
 $(L_{p'} \leq TLV) \& \& (R_{p'} \leq \mu_{p'} \cdot (1 - \pi_{p'}^B))$;
- 15 **if** $\sum_{p \in \mathcal{P}} U_p$ can be improved with variation ΔR_p **then**
- 16 $R_p = R_p + \Delta R_p$, $R_{p'} \leftarrow R_{p'} - \Delta R_p$;
- 17 Update the available resources of path p and p' ;
- 18 **end**
- 19 **end**

ratio of the proposed algorithm is mainly in the range of [0.7, 0.9]. As depicted in Fig. 4(b), the proposed algorithm achieves performance values closer to the optimum while concurrently using cellular and Wi-Fi networks. Higher variations in the approximation ratio are observed while using three access networks.

Due to the complexity of the considered problem, the theoretical approximation ratio of the proposed algorithm is not provided in this paper. In real video streaming systems, the energy consumption and video quality are affected by many factors, e.g., time-varying channel status, user mobility, and video content variation. Therefore, the theoretical bounds provided by mathematical derivations may vary significantly with experimental/evaluation results.

C. Retransmission Control

The purpose of the retransmission control algorithm in DEAM is to reduce bandwidth usage and device energy consumption. In the conventional retransmission algorithms of MPTCP, the mobile energy consumption and video delay constraint are not considered. Such delay and energy agnostic retransmission may incur unnecessary energy waste and degrade video quality. Each communication path is analyzed with the following parameters: round trip time (RTT_p), congestion window size ($cwnd_p$), slow start threshold ($ssthresh_p$) and retransmission timeout (RTO_p). DEAM provides aggregate-level ACKs (acknowledgements) of received packets from all communication paths to the sender side. The latest aggregate feedback is sent by the receiver side immediately after receiving each packet. In particular, the ACK packets in the proposed retransmission

algorithm are transmitted through the most reliable uplink channel to the sender side. This approach mitigates the problem of dropped/overdue feedback packets and increases the measurement accuracy.

If an ACK packet is received from any communication path, the packet loss rates π_p^B are updated. We update the round-trip time RTT_p for each communication path with the equation

$$RTT_p = \begin{cases} \frac{cwnd_p}{\mu_p}, & \text{if } \mu_p \cdot \tau_p < cwnd_p, \\ \tau_p + \frac{MTU}{\mu_p}, & \text{if } \mu_p \cdot \tau_p \geq cwnd_p, \end{cases}$$

where τ_p represents the propagation delay of the communication path p . This propagation latency can be estimated based on the sequence numbers and timestamps of successive received packets.

The different physical characteristics of heterogeneous access networks result in path asymmetry in the network status. The uniform congestion control will most likely induce performance degradations due to the low-quality communication paths. The aggregate selective ACK feedback is filtered to check the delivery status of the transmitted packets for a specific communication path. The ACK packet is adopted to the congestion window of each path. A retransmission timeout (RTO_p) is set while delivering a packet to path p and the timeout value is calculated with the equation $RTO_p = RTT_p + 4 \times \sigma_{RTT_p}$.

The response to a timeout is identical to the conventional TCP. DEAM does not perform fast retransmissions since it may result in unnecessary transmissions. The slow start threshold ($ssthresh_p$) will be set to $\max(cwnd_p/2, 4 \times MTU)$ after receiving four duplicated selective acknowledgments.

The multipath congestion control algorithm in DEAM is also regulated to achieve TCP friendliness. We assume $\Psi(cwnd_p)$ and $\Upsilon(cwnd_p)$ represent the increasing and decreasing functions of the congestion window size for path p , respectively. The sufficient and necessary condition to achieve TCP-friendly congestion control is $\Psi(cwnd_p) = \frac{3 \times \Upsilon(cwnd_p)}{2 - \Upsilon(cwnd_p)}$ [52]. In the emulations conducted in this work, we set the congestion window adaption with the equations $\Psi(cwnd_p) = \frac{3 \times \theta}{2 \times \sqrt{cwnd_p + 1 - \theta}}$, $\Upsilon(cwnd_p) = \frac{\theta}{\sqrt{cwnd_p + 1}}$, where θ is selected from the values of {0.1, 0.2, ..., 0.9}. We set the value of θ to be 0.5 in the performance evaluation [the same as the AIMD (additive increase multiplicative decrease) algorithm in TCP].

The energy consumption and delay constraints are not considered in existing MPTCP retransmission schemes. This paper proposes an energy- and delay-aware retransmission control scheme to prevent unnecessary retransmissions for conserving mobile energy and network bandwidth. Specifically, the timeout packet will be retransmitted through the path with lowest energy consumption to aim at enforcing the deadline-constrained packet delivery. In the context of heterogeneous access medium, the packet losses may be caused by the wireless losses (e.g., external interference and channel fading) and network congestions (e.g., buffer overflow and link fault). The conventional MPTCP

Algorithm 3: Energy and Delay Aware Packet Retransmission Control

Input: $\{RTT_p, \mu_p, \pi_p^B, m_p\}_{p \in \mathcal{P}}$, $\mathcal{R} = \{R_p\}_{p \in \mathcal{P}}$, T, e_p ;
Output: $\{ssthresh_p, cwnd_p\}_{p \in \mathcal{P}}$, packet retransmission path;

- 1 $\overline{RTT}_p = \frac{1}{16} \times RTT_p + \frac{15}{16} \times \overline{RTT}_p$
- $\sigma_{RTT_p} = \frac{1}{8} \times |RTT_p - \overline{RTT}_p| + \frac{7}{8} \times \sigma_{RTT_p}$;
- 2 $C_{\text{I}} = (RTT_p < \overline{RTT}_p - \sigma_{RTT_p}) \&\& (m_p == 1)$;
- $C_{\text{II}} = (RTT_p < \overline{RTT}_p - \sigma_{RTT_p}/2) \&\& (m_p == 2)$;
- $C_{\text{III}} = (RTT_p < \overline{RTT}_p) \&\& (m_p == 3)$;
- $C_{\text{IV}} = (RTT_p < \overline{RTT}_p - \sigma_{RTT_p}/2) \&\& (m_p > 3)$;
- 3 $\text{Rand_Loss_Cond} = C_{\text{I}} || C_{\text{II}} || C_{\text{III}} || C_{\text{IV}}$;
- 4 **if** the number of received packets is less than the number of sent packets in path p **then**
- 5 **if** the sender receives three ACK packets with the same sequence number **then**
- 6 $ssthresh_p = \max(cwnd_p/2, 4 \times MTU)$;
- 7 $cwnd_p = ssthresh_p$;
- 8 **end**
- 9 **Random packet loss identification:** **if**
- 10 $\text{Rand_Loss_Cond} == \text{true}$ **then**
- 11 $cwnd_p = MTU$ \triangleright Random packet loss occurs;
- 12 $ssthresh_p = \max(cwnd_p/2, 4 \times MTU)$;
- 13 **end**
- 14 **Delay and energy aware retransmission:**
- 15 Select the path set \mathcal{P}' with the delay conditions of $D_{p \in \mathcal{P}'} < T$;
- 16 $p_{\min} = \arg \min_{p \in \mathcal{P}'} \{e_p\}$;
- 17 Retransmit the lost packet through p_{\min} ;
- 18 **end**

retransmission does not include mechanisms for the packet loss differentiation. In the design of DEAM, the packet loss differentiation is based on the average round-trip time \overline{RTT} , standard deviations σ_{RTT_p} , and number of consecutive losses m_p [53]. The details of the delay- and energy-aware retransmission control scheme are presented in Algorithm 3.

V. PERFORMANCE EVALUATION

This section evaluates the performance of the proposed DEAM by conducting extensive semi-physical emulations based on the Exata platform [54]. The first subsection describes the evaluation methodology that includes the setup, metrics, competing schemes, and experimental scenarios. Second, the evaluation results are presented and discussed in detail. The componentwise validation of the proposed DEAM is provided in the supplementary file.

In summary, the performance improvements of the proposed DEAM solution are listed as follows.

- 1) DEAM reduces the energy consumption by up to 62.3 (24.3%), 86.8 (32.3%) and 125.6 (45.6%) J (Joule) compared to ADMIT [34], EMTCP [30] and MPTCP [16] for 200 seconds, respectively.
- 2) DEAM improves the video PSNR by up to 5.8 (23.2%), 7.3 (28.6%) and 9.8 (39.3%) dB compared to ADMIT, EMTCP and MPTCP with the same energy consumption, respectively.

- 3) DEAM reduces the buffering rate by up to 4.3 (23.2%), 6.2 (28.9%) and 13.2 (37.7%) times/min compared to ADMIT, EMTCP and MPTCP, respectively.
- 4) DEAM increases the number of effective retransmissions by up to 15.8 (35.2%), 22.3 (46.3%) and 36.7 (58.2%) compared to ADMIT, EMTCP and MPTCP, respectively.

A. Evaluation Methodology

Emulation Setup : JM and Exata are adopted as the video codec and network emulator, respectively. Fig. 5 presents the system architecture of the performance evaluation. The main emulation configurations are set as follows.

Network Emulator: In this evaluation framework, Exata 2.1 [54] (i.e., an advanced edition of QualNet [55]) is adopted as the network emulator to perform semi-physical emulations. To conduct the evaluations with real-time H.264 videos, the source code of JM is integrated with Exata to develop an application for video data transfer. The Exata Programmers Guide [54] presents the detailed descriptions of the development steps. In the emulation topology, the mobile terminal is equipped with three wireless network interfaces, i.e., cellular, Wi-Fi and WiMAX. Table III summarizes the parameter configurations of different wireless networks [3], [36]. As MPTCP is not included at the transport-layer stack of the Exata emulator, the transport protocol is developed on the basis of the C code from GitHub [56] and the standard TCP module. The Linux kernel version is 14.04 and the MPTCP version is v0.89 [57].

As illustrated in Fig. 5, each router in the core networks is connected to an edge node that injects background traffic. Each of the edge nodes connects to four background traffic generators introducing cross-traffic with a Pareto distribution. Based on real traffic traces collected on the Internet, the packet sizes of background traffic are varied as follows: 43% packets are large (≥ 1400 Bytes), 17% packets are medium [(144, 1400) Bytes], and 40% packets are small (≤ 144 Bytes) [58]. The overall cross-traffic loads imposed on the communication paths are configured to vary randomly from 20 to 40 percent of the bottleneck links bandwidth. The data scheduling cycle is configured as 500 ms (the playback duration of a GoP), and the TLV (Threshold Limit Value) is 1.2 [3], [47]. The packet interleaving level (ω_p) is configured as 5 ms for each communication link to the mobile terminal.

Video Codec: This research adopts the H.264/AVC reference software JM 18.2 [59] as the video codec. The source videos are compressed at 30 frames per second and a GoP includes 15 frames. The GoP structure is IPP...P to enable low-delay video coding. The video test sequences are river bed, blue sky, park joy, and mobcal in HD (high-definition) format. These test sequences feature different spatial features and temporal motions in rate-distortion profile. To obtain statistically meaningful results, we concatenate the video sequences to be 6000 frame-long. The delay constraint (T) for each GoP is set to be 500 ms to prevent playback buffer starvation. The streaming video is encoded at the bit rates of 2.4, 2.2, 2.8, and 1.85 Mbps for the mobile trajectory indexed from I to IV. These trajectories

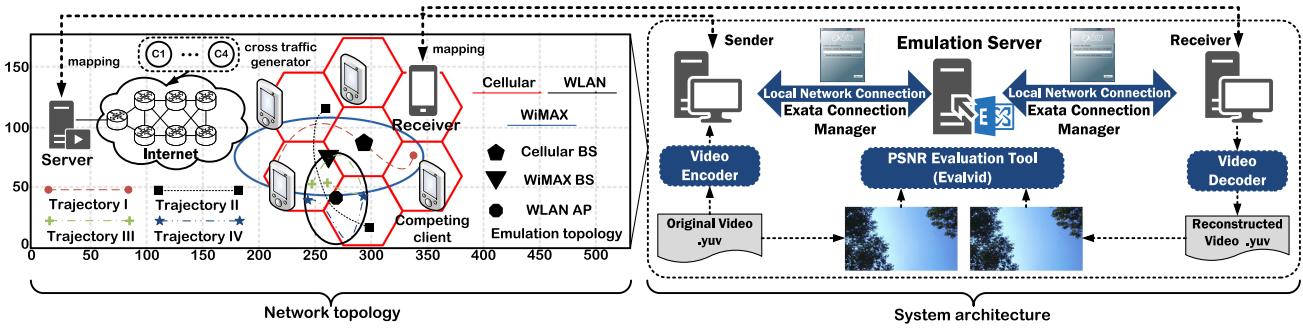


Fig. 5. System architecture and network topology for performance evaluation.

TABLE III
CONFIGURATIONS OF WIRELESS NETWORKS

| WLAN parameter | Value |
|-------------------------------------|----------------------|
| Average channel bit rate | 2 Mbps |
| Slot time | 10 μ s |
| Maximum contention window | 32 |
| $\mu_p, \pi_p^B, 1/\xi_p^B$ | 1000 Kbps, 6%, 12 ms |
| Cellular parameter | Value |
| Common control channel power | 33 dB |
| Maximum power of BS | 43 dB |
| Total cell bandwidth | 3.84 Mc/s |
| Target SIR value | 10 dB |
| Orthogonality factor | 0.4 |
| Inter/intra cell interference ratio | 0.55 |
| Background noise power | -106 dB |
| $\mu_p, \pi_p^B, 1/\xi_p^B$ | 1500 Kbps, 2%, 10 ms |
| WiMAX parameter | Value |
| System bandwidth | 7 MHz |
| Number of carriers | 256 |
| Sampling factor | 8/7 |
| Average SNR | 15 dB |
| Symbol duration | 2048 |
| $\mu_p, \pi_p^B, 1/\xi_p^B$ | 1200 Kbps, 4%, 15 ms |

TABLE IV
AVERAGE DISTORTION PARAMETERS FOR VIDEO SEQUENCES

| Video sequence | α | β |
|----------------|--------------------|---------|
| blue sky | 1.78×10^5 | 1739 |
| mobcal | 1.45×10^5 | 1262 |
| park joy | 1.83×10^4 | 531 |
| river bed | 1.62×10^4 | 347 |

indicate different access options and capacity constraints for the mobile user.

Performance Metrics :

- Mobile energy and power. The mobile energy consumption is measured in Joules (J) and the power dissipation is expressed in units of milliWatts (mW). These values are measured using the energy module in Exata 2.1 [54]. We have also measured the real-time power consumption using the Monsoon monitor and the receiving traces in mobile trajectories I and III.
- PSNR. In both research and industrial communities, the PSNR (peak signal-to-noise ratio) [60] is a widely-adopted

metric to quantify video quality. This parameter is calculated as a function of the MSE (mean squared error) between the original and reconstructed video frames [60]. If a video frame experiences either transmission or overdue loss, this frame is considered as invalid. Consequently, the concealment method will copy the payload from the last received frame before the lost one.

- Retransmissions. The efficacy of packet retransmission mechanism in MPTCP is critical to conserve network bandwidth and device energy. The number of total and effective retransmissions are measured for comparison.
- Buffering rate [61]. This metric captures the frequency at which buffering events occur during the session and is computed as the ratio of the number of buffering events to the duration of the session.

Reference Schemes :

- ADMIT [34]. The quality-driven MPTCP adapts the FEC coding and subflow allocation based on utility maximization. The data packets are distributed onto the communication paths based on reliability value.
- EMTCP [30]. In the energy-efficient MPTCP scheme, the authors take advantage of the throughput-energy tradeoff in path selection and congestion control. The algorithms for real-time applications in [24] are implemented in the emulations for comparison.
- MPTCP [17]. We also adopt the baseline MPTCP as a reference in the performance evaluation. The path selection, congestion window adaptation and packet retransmission mechanisms are specified in the RFC 6182 [16] and 6824 [17].

In addition to the above schemes, we also present the optimal results obtained with the *oracle* for comparison.

Emulation Scenarios : The semi-physical emulations are performed in mobile scenarios and Fig. 5 depicts the trajectories. To validate the effectiveness in mobile energy conservation of all the competing schemes, the target video quality is configured as 25, 31 and 37 dB. As defined in ITU-T P. 910 [62], these PSNR values represent the acceptable, good and excellent quality levels.

B. Evaluation Results

Energy and Power Consumption : Fig. 6(a) presents the average mobile energy consumption of all the evaluated schemes

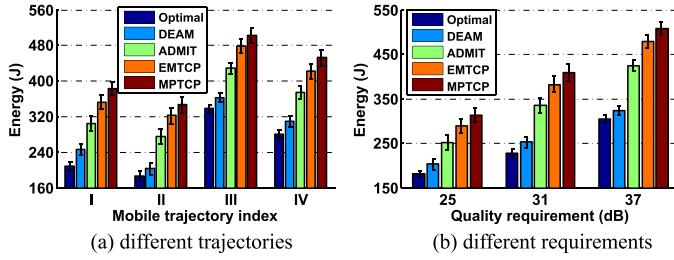


Fig. 6. Comparison of energy consumption.

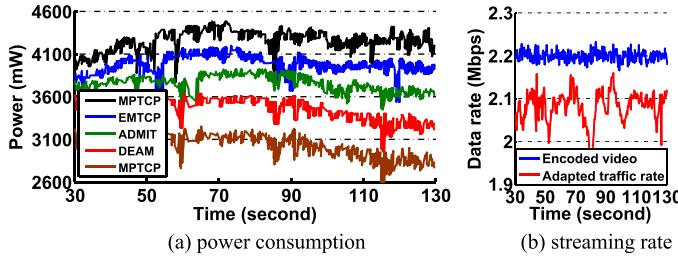


Fig. 7. Power consumption and video streaming rate during the interval of [30, 130] seconds.

in different mobile trajectories. The proposed DEAM achieves the lower mobile energy consumption in all the emulation scenarios lasting for 200 seconds. DEAM outperforms the ADMIT and EMTCP as we consider both the mobile energy and end-to-end distortion in the subflow allocation. With regard to the special feature of video streaming, a higher data transmission throughput cannot always reduce the video channel distortion. For instance, allocating subflow to a communication path with available bandwidth and large propagation delay will improve the throughput but also degrades the average video quality. This is because there will be substantial amount of overdue packets on this communication path.

The mobile energy for different video quality constraints along the Trajectory I is plotted in Fig. 6(b). As observed from the measurement results, the superiority level of DEAM over the competing MPTCP solutions becomes higher as the quality requirement increases. The tradeoff between the received video quality and the mobile energy consumption is also verified by the results.

The mobile power dissipation of the evaluated MPTCP solutions during the interval of [30, 130] seconds is presented in Fig. 7(a). The results are measured with the energy module in Exata emulation platform [54]. DEAM achieves lower power values and variations during the multipath video session. As shown in Fig. 7(b), DEAM reduces the video transmission rate with regard to the imposed video quality and thus minimizes the energy consumption for data transfer. In particular, the fluctuations of the original video traffic rate in Fig. 7(b) are caused by the bit rate variability [63]. The allocation rates (adjusted video traffic rates) of the baseline MPTCP, EMTCP and ADMIT schemes are the same as the original video traffic rate, since these schemes do not proactively drop video frames for rate adaptation.

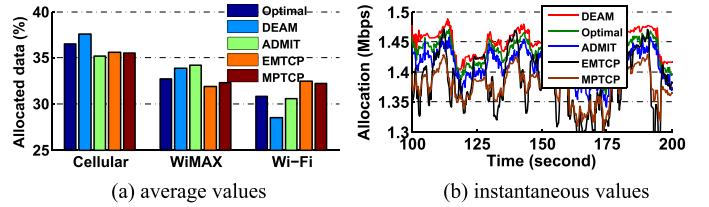


Fig. 8. Rate allocation over different access networks.

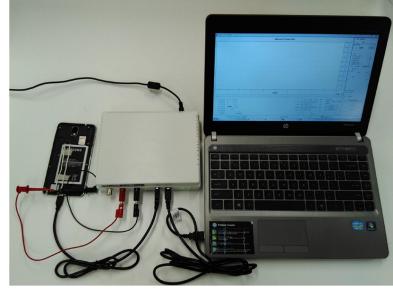


Fig. 9. Measuring power consumption with the Monsoon monitor.

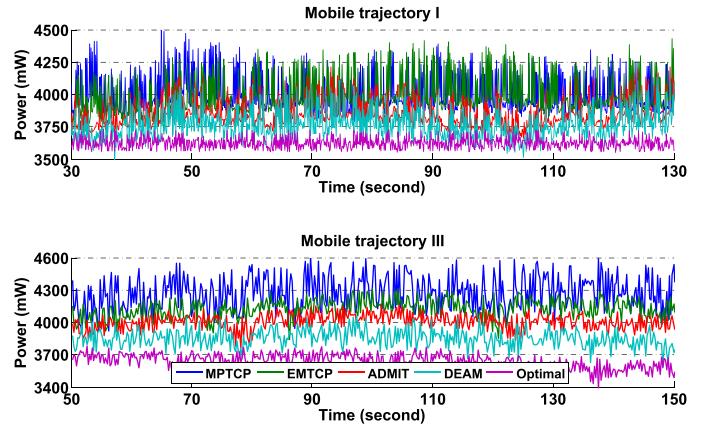


Fig. 10. Real-time power consumption during the intervals of [30, 130] (Trajectory I) and [50, 150] (Trajectory III) seconds.

Fig. 8 presents the rate allocation for different access networks for all the evaluated schemes. In general, the percentage of allocated data is proportional to the communication path quality. The reference schemes assign more video data than the proposed DEAM to the Wi-Fi.

Specifically, the mobile power consumption is measured with the Monsoon monitor and received video traffic trace (of mobile trajectories I and III) as shown in Fig. 9. The instantaneous values of power consumption during the intervals of [30, 130] (Trajectory I) and [50, 150] (Trajectory III) seconds are shown in Fig. 10. The result is measured based on the mobile power consumption value for downloading the traffic traces with LTE and Wi-Fi interfaces separately. These values are different from those shown in Fig. 7(a) because: 1) the power consumption values in Fig. 7(a) are obtained with the energy module in Exata [54] while the values in Fig. 8 are measured with the Monsoon monitor and Samsung smartphone based on the received traffic trace and 2) the real-time power consumption of mobile

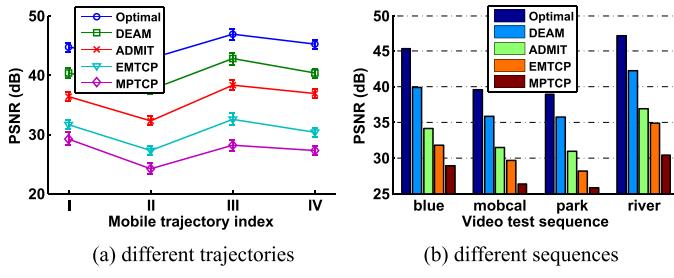


Fig. 11. Comparison of average PSNR results.

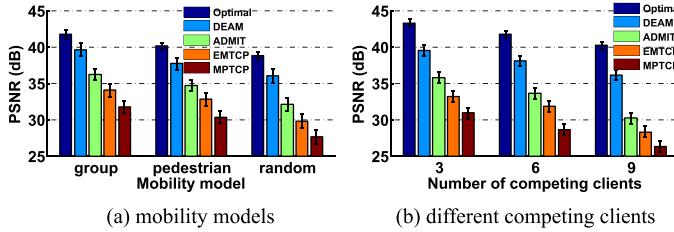


Fig. 12. Comparison of average PSNR results in different mobility models and competing clients.

devices is affected by many random factors, e.g., random channel behavior, video bitrate variability, external interference, etc. Even in the same emulation scenario, there may be significantly different power drops and rises. DEAM achieves lower power energy consumption as it employs the video traffic load adaptation and chooses the path to minimize the mobile energy consumption.

PSNR: The quality requirement (\bar{D}) of the proposed DEAM is gradually adjusted to achieve the same energy consumption level as the reference MPTCP schemes. Fig. 11(a) shows the measured average PSNR results in different mobile trajectories and we can observe the results' pattern is almost the same as shown in Fig. 6(a). It is also in accordance with the Proposition 1, i.e., higher video quality generally requires higher energy consumption. DEAM outperforms the reference schemes and the performance gaps become larger in Trajectory III, which indicates the superiority of DEAM in overcoming the path diversity in heterogeneous access networks and integrating the available bandwidth. Fig. 11(b) shows the average PSNR values measure from different video test sequences in mobile Trajectory III.

Fig. 12 shows the average PSNR values for different mobility models. The results indicate all the evaluated schemes achieve the highest PSNR values in the pedestrian mobility, and the lowest video quality under the random waypoint. We can also observe the largest performance gaps between the proposed DEAM and the reference schemes in the random waypoint model. In this random-based mobility model, the location, velocity and acceleration of the mobile clients are chosen randomly and independently. However, the group and pedestrian models have dependencies and restrictions.

Fig. 13 compares the instantaneous PSNR values for the video frames numbered from 1500 to 2000. These microscopic values are measured from the blue sky test sequence while the mobile is moving along Trajectory IV. DEAM delivers video streams with higher PSNR values with lower fluctuations to achieve

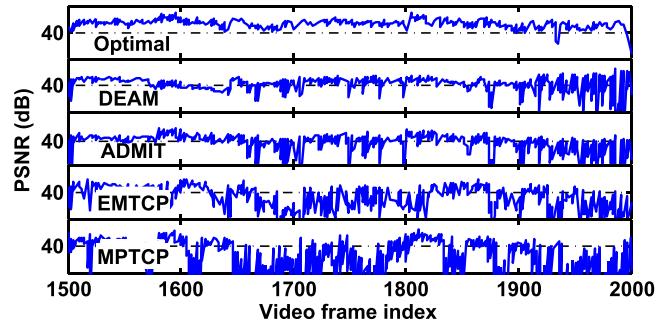


Fig. 13. Comparison of PSNR per video frame indexed from 1500 to 2000.

the excellent perceptual quality (above 37 dB). However, the reference schemes (ADMIT, EMTCP and baseline MPTCP) frequently incur violations of imposed distortion constraint.

Fig. 14 presents the representative values of the received video frames with the evaluated MPTCP solutions. The perceptual quality results are compared with 1375-th frame from the mobcal sequence and the 1187-th frames of the blue sky sequence. A clear image is received with the proposed DEAM, while more transmission impairments are observed on the video frames delivered with ADMIT, EMTCP and MPTCP.

We plot the instantaneous PSNR values for the video frames indexed from 1500 to 2000 in Fig. 13 to compare the microscopic quality results. The instantaneous values are measured from the *blue sky* sequence while the client moving along Trajectory IV. DEAM achieves higher values with lower variations to guarantee the excellent perceived quality (above 37 dB) while the reference schemes frequently induce violations of imposed distortion constraint.

The typical results of the received video frames with the competing MPTCP schemes are presented in Fig. 14. The subjective quality results are measured from the 1187-th frame of the *blue sky* sequence and 1375-th frame from the *mobcal* sequence. A clear image is obtained with the proposed DEAM, while more impairments are observed on the video frames received with ADMIT, EMTCP and MPTCP.

Retransmissions: The total and effective packet retransmission numbers of the evaluated MPTCP solutions are plotted in Fig. 15(a). DEAM achieves higher effective retransmission number with fewer total packet retrasmssions as the retransmission control scheme includes the delay constraint of video applications. Moreover, the retransmission packets are delivered through the communication path to achieve low energy consumption and on-time delivery. A higher effective retransmission number also enhances the average goodput as depicted in Fig. 15(b).

Out-of-order packets also significantly impact on perceived video quality and mobile energy consumption of multipath video transmission. These out-of-order arrival problems are caused by the path heterogeneity and incur larger end-to-end transmission latency of the video frames. The out-of-packet number during the interval of [50, 100] seconds is profiled in Fig. 16. DEAM guarantees high in-order arrival probability of packets allocated onto different paths by overcoming the asymmetry in heterogeneous networks.

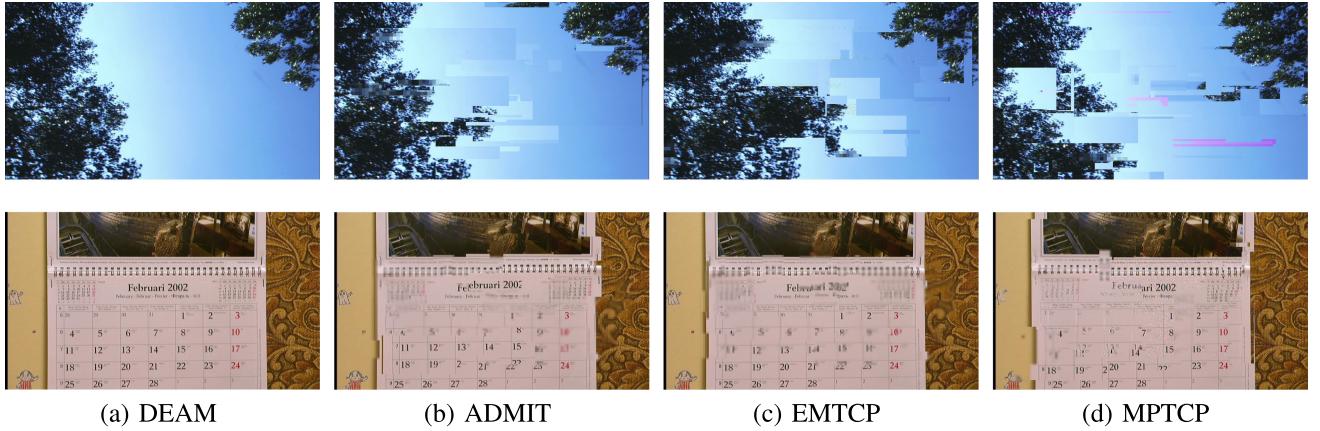


Fig. 14. Comparison of subjective video quality measured from the *blue sky* and *mobcal* sequences.

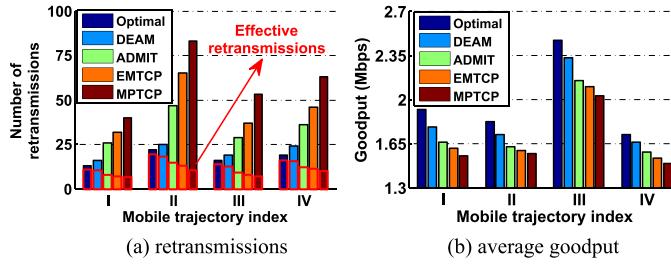


Fig. 15. Comparison of retransmission and goodput performance.

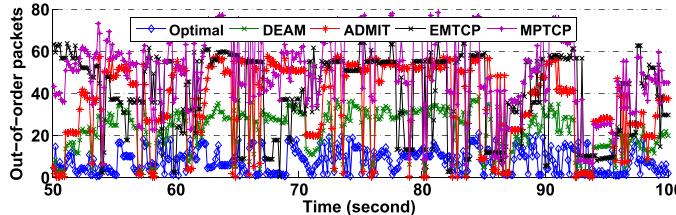


Fig. 16. Evolutions for the number of out-of-order packets during the interval [50, 100] seconds.

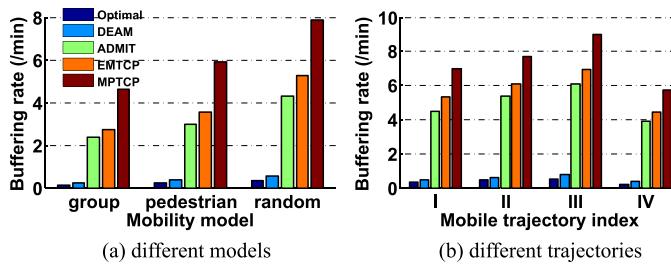


Fig. 17. Comparison of average buffering rate.

Fig. 17 presents the average buffering rates in different mobility models and trajectories. The results' pattern is almost opposite to that shown Fig. 11(a) and Fig. 12(a). A higher buffering rate generally leads to lower perceived quality [63]. The lower buffering ratio of DEAM is achieved by the delay performance superiority to prevent buffer starvation.

VI. CONCLUSION AND DISCUSSION

Mobile energy consumption is critical to sustain power-intensive mobile video streaming with delay and quality requirements. This research presents a Delay-Energy-quAlity aware MPTCP (DEAM) scheme to stream energy-efficient and quality-guaranteed videos over heterogeneous wireless networks. Compared with the previous studies on MPTCP, we advance the state of the art by introducing the delay-constrained energy-quality tradeoff into the decision process of concurrent multipath transmission. This paper develops solutions for the subflow allocation and retransmission control, which significantly conserve the mobile energy for radio communications while respecting quality constraint. The emulation results show that DEAM exceeds the reference MPTCP solutions in mobile energy conservation and user-perceived quality. As future work, we will consider improving the congestion control [64], [65] and send buffer management algorithms in DEAM to further improve video data throughput.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the anonymous reviewers and Associate Editor for the valuable comments.

REFERENCES

- [1] Download Booster, 2015. [Online]. Available: <http://galaxys5guide.com/samsung-galaxy-s5-features-explained/galaxy-s5-download-booster/>
- [2] Mushroom products, 2016. [Online]. Available: <https://www.mushroomnetworks.com/>
- [3] J. Wu, B. Cheng, C. Yuen, Y. Shang, and J. Chen, "Distortion-aware concurrent multipath transfer for mobile video streaming in heterogeneous wireless networks," *IEEE Trans. Mobile Comput.*, vol. 14, no. 4, pp. 688–701, Apr. 2015.
- [4] D. Ho, G. S. Park, and H. Song, "Game-theoretic scalable offloading for video streaming services over LTE and WiFi networks," *IEEE Trans. Mobile Comput.*, vol. 17, no. 5, pp. 1090–1104, May 2018.
- [5] M. Xing, S. Xiang, and L. Cai, "A real-time adaptive algorithm for video streaming over multiple wireless access networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 4, pp. 795–805, Apr. 2014.
- [6] J. Wu, C. Yuen, B. Cheng, Y. Shang, and J. Chen, "Goodput-aware load distribution for real-time traffic over multipath networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 8, pp. 2286–2299, Aug. 2015.
- [7] Cisco, "Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021," San Jose, CA, USA, Feb. 2017.

- [8] C. Xu, Z. Li, J. Li, H. Zhang, and G.-M. Muntean, "Cross-Layer fairness-driven concurrent multipath video delivery over heterogeneous wireless networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 7, pp. 1175–1189, Jul. 2015.
- [9] J. Wu, J. Yang, X. Wu, and J. Chen, "A low latency scheduling approach for high definition video streaming over heterogeneous wireless networks," in *Proc. IEEE GLOBECOM*, 2013, pp. 1723–1729.
- [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 Comput.*, vol. 12, no. 9, pp. 1737–1749, Sep. 2013.
- [11] J. Wu *et al.*, "Loss tolerant bandwidth aggregation for multihomed video streaming over heterogeneous wireless networks," *Wireless Pers. Commun.*, vol. 74, no. 2, pp. 1265–1282, 2014.
- [12] N. Freris, C. Hsu, J. Singh, and X. Zhu, "Distortion-aware scalable video streaming to multinetwork clients," *IEEE/ACM Trans. Netw.*, vol. 21, no. 2, pp. 469–481, Apr. 2013.
- [13] V. Sharma *et al.*, "A transport protocol to exploit multipath diversity in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 20, no. 4, pp. 1024–1039, Aug. 2012.
- [14] K. Piamrat, A. Ksentini, J. M. Bonnin, and C. Viho, "Radio resource management in emerging heterogeneous wireless networks," *Comput. Commun.*, vol. 34, no. 9, pp. 1066–1076, 2011.
- [15] J. Yoon, H. Zhang, S. Banerjee, and S. Rangarajan, "MuVi: A multicast video delivery scheme for 4G cellular networks," in *Proc. ACM MobiCom*, 2012, pp. 209–220.
- [16] A. Ford *et al.*, "Architectural guidelines for multipath TCP development," Internet Engineering Task Force, Fremont, CA, USA, RFC 6182, 2011.
- [17] A. Ford *et al.*, "TCP extensions for multipath operation with multiple addresses," Internet Engineering Task Force, Fremont, CA, USA, RFC 6824, 2013.
- [18] R. Stewart, "Stream Control Transmission Protocol (SCTP)," Internet Engineering Task Force, Fremont, CA, USA, RFC 4960, 2007.
- [19] V. Singh, S. Ahsan, and J. Ott, "MPRTP: Multipath considerations for real-time media," in *Proc. ACM 4th Int. Conf. Multimedia Syst.*, 2013, pp. 190–201.
- [20] S. Floyd and K. Fall, "Promoting the use of end-to-end congestion control in the Internet," *IEEE/ACM Trans. Netw.*, vol. 7, no. 4, pp. 458–472, Aug. 1999.
- [21] Y. C. Chen *et al.*, "A Measurement-based study of multipath TCP performance over wireless networks," in *Proc. ACM Internet Meas. Conf.*, 2013, pp. 455–468.
- [22] D. Wischik *et al.*, "Design, implementation and evaluation of congestion control for multipath TCP," in *Proc. USENIX Symp. Netw. Syst. Design Implementation*, 2011, pp. 99–112.
- [23] C. Raiciu *et al.*, "How hard can it be? Designing and implementing a deployable multipath TCP," in *Proc. USENIX Symp. Netw. Syst. Design Implementation*, 2012, p. 29.
- [24] B. Hesmans, G. Detal, R. Bauduin, and O. Bonaventure, "SMAPP: Towards smart multipath TCP-enabled APplications," in *Proc. ACM 10th Int. Conf. Emerg. Netw. Exp. Technol.*, 2015, Art. no. 28.
- [25] S. Deng, R. Netravali, A. Sivaraman, and H. Balakrishnan, "WiFi, LTE, or Both? Measuring multi-homed wireless internet performance," in *Proc. ACM Internet Meas. Conf.*, 2014, pp. 181–194.
- [26] M. A. Hoque, M. Siekkinen, and J. K. Nurminen, "Using crowd-sourced viewing statistics to save energy in wireless video streaming," in *Proc. ACM MobiCom*, 2013, pp. 377–388.
- [27] M. A. Hoque, M. Siekkinen, and J. Nurminen, "Energy efficient multimedia streaming to mobile devices: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 579–597, Jan./Mar. 2014.
- [28] M. Ismail and W. Zhuang, "Mobile terminal energy management for sustainable multi-homing video transmission," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4616–4627, Aug. 2014.
- [29] M. Ismail, W. Zhuang, and S. Elhedi, "Energy and content aware multi-homing video transmission in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3600–3610, Jul. 2013.
- [30] Q. Peng, M. Chen, A. Walid, and S. H. Low, "Energy-efficient multipath TCP for mobile devices," in *Proc. ACM MobiHoc*, 2014, pp. 257–266.
- [31] Y. Cui, L. Wang, X. Wang, H. Wang, and Y. Wang, "FMTCP: A fountain code-based multipath transmission control protocol," in *Proc. IEEE 32th Int. Conf. Distrib. Comput. Syst.*, 2012, pp. 366–375.
- [32] M. Li *et al.*, "Tolerating path heterogeneity in multipath TCP with bounded receive buffers," *Comput. Netw.*, vol. 64, pp. 1–14, 2013.
- [33] Y. Lim *et al.*, "Cross-layer path management in multi-path transport protocol for mobile devices," in *Proc. IEEE INFOCOM*, 2014, pp. 1815–1823.
- [34] J. Wu, C. Yuen, B. Cheng, M. Wang, and J. Chen, "Streaming high-quality mobile video with multipath TCP in heterogeneous wireless networks," *IEEE Trans. Mobile Comput.*, vol. 15, no. 9, pp. 2345–2361, Sep. 2016.
- [35] W. Song and W. Zhuang, "Performance analysis of probabilistic multipath transmission of video streaming traffic over multi-radio wireless devices," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1554–1564, Apr. 2012.
- [36] J. Wu *et al.*, "Joint source-channel coding and optimization for mobile video streaming in heterogeneous wireless networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2013, 2013, Art. no. 283.
- [37] W. Hu and G. Cao, "Energy-aware video streaming on smartphones," in *Proc. IEEE INFOCOM*, 2015, pp. 1185–1193.
- [38] D. H. Bui *et al.*, "Greenbag: Energy-efficient bandwidth aggregation for real-time streaming in heterogeneous mobile wireless networks," in *Proc. IEEE Real-Time Syst. Symp.*, 2013, pp. 57–67.
- [39] H. Han, S. Shakottai, C. Hollot, R. Srikanth, and D. Towsley, "Multipath TCP: A joint congestion control and routing scheme to exploit path diversity in the Internet," *IEEE/ACM Trans. Netw.*, vol. 14, no. 6, pp. 1260–1271, Dec. 2006.
- [40] K. Stuhlmüller, N. Färber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 1012–1032, Jun. 2000.
- [41] E. Harjula, O. Kassinen, and M. Ylanttila, "Energy consumption model for mobile devices in 3G and WLAN networks," in *Proc. IEEE Consum. Commun. Netw. Conf.*, 2012, pp. 532–537.
- [42] J. Huang, F. Qian, A. Gerber, *et al.*, "A close examination of performance and power characteristics of 4G LTE networks," in *Proc. ACM Mobicom*, 2012, pp. 225–238.
- [43] E. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, no. 9, pp. 1253–1265, 1960.
- [44] G. Hasslinger, A. Schwahn, and F. Hartleb, "2-State (semi-) Markov processes beyond Gilbert-Elliott: Traffic and channel models based on 2nd order statistics," in *Proc. IEEE INFOCOM*, 2013, pp. 1438–1446.
- [45] J. Wu, B. Cheng, C. Yuen, N.-M. Cheung, and J. Chen, "Trading delay for distortion in one-way video communication over the Internet," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 4, pp. 711–723, Apr. 2016.
- [46] L. Zhou *et al.*, "Distributed scheduling scheme for video streaming over multi-channel multi-radio multi-hop wireless networks," *IEEE J. Sel. Area Commun.*, vol. 28, no. 3, pp. 409–419, Apr. 2010.
- [47] X. Zhu *et al.*, "Distributed rate allocation policies for multihomed video streaming over heterogeneous access networks," *IEEE Trans. Multimedia*, vol. 11, no. 4, pp. 752–764, Jun. 2009.
- [48] T. Oliveira, S. Mahadevan, and D. P. Agrawal, "Handling network uncertainty in heterogeneous wireless networks," in *Proc. IEEE INFOCOM*, 2011, pp. 2390–2398.
- [49] M. R. Garey and D. S. Johnson, *Computers and Intractability*. San Francisco, CA, USA: Freeman, 2002.
- [50] S. Wenger and T. Stockhammer, "RTP payload format for H. 264 video," Internet Engineering Task Force, Fremont, CA, USA, RFC 3984, 2005.
- [51] D. Jurca and P. Frossard, "Video packet selection and scheduling for multipath streaming," *IEEE Trans. Multimedia*, vol. 9, no. 3, pp. 629–641, Apr. 2007.
- [52] L. Cai, X. Shen, J. Pan, and J. Mark, "Performance analysis of TCP-friendly AIMD algorithms for multimedia applications," *IEEE Trans. Multimedia*, vol. 7, no. 2, pp. 339–355, Apr. 2005.
- [53] S. Cen, P. C. Cosman, and G. M. Voelker, "End-to-end differentiation of congestion and wireless losses," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 703–717, Oct. 2003.
- [54] Exata, 2013. [Online]. Available: <http://www.scalable-networks.com/exata>
- [55] QualNet, 2013. [Online]. Available: <http://www.scalable-networks.com/qualnet>
- [56] MPTCP GitHub repository, 2016. [Online]. Available: <https://github.com/multipath-tcp/mptcp/>
- [57] C. Paasch *et al.*, Multipath TCP in the Linux Kernel, 2016. [Online]. Available: <http://multipath-tcp.org/pmwiki.php/Researchers/References>
- [58] P. Borgnat *et al.*, "Seven years and one day: Sketching the evolution of internet traffic," in *Proc. IEEE INFOCOM*, 2009, pp. 711–719.
- [59] JM Homepage, 2004. [Online]. Available: <http://iphone.hhi.de/suehring/tm/>
- [60] ANSI T1.TR.74-2001, "Objective video quality measurement using a peak-signal-to-noise-ratio (PSNR) full reference technique," 2001. [Online]. Available: <http://webstore.ansi.org/RecordDetail.aspx?sku=T1.TR.74-2001>

- [61] A. Balachandran *et al.*, “Developing a predictive model of quality of experience for internet video,” in *Proc. ACM SIGCOMM*, 2013, pp. 339–350.
- [62] *Subjective Video Quality Assessment Methods for Multimedia Application*, ITU-T Recommendation P. 910, Apr. 2008.
- [63] G. Auwera and M. Reisslein, “Implications of smoothing on statistical multiplexing of H. 264/AVC and SVC video streams,” *IEEE Trans. Broadcast*, vol. 55, no. 3, pp. 541–558, Sep. 2009.
- [64] C. Raiciu *et al.*, “Coupled Congestion Control for Multipath Transport Protocols,” Internet Engineering Task Force, Fremont, CA, USA, RFC 6356, 2011.
- [65] R. Khalili *et al.*, “MPTCP is not pareto-optimal: Performance issues and a possible solution,” in *Proc. ACM 8th Int. Conf. Emerg. Netw. Exp. Technol.*, 2012, pp. 1–12.
- [66] J. Wu, B. Cheng, and M. Wang, “Energy minimization for quality-constrained video with multipath TCP over heterogeneous wireless networks,” in *Proc. IEEE 36th Int. Conf. Distrib. Comput. Syst.*, 2016, pp. 487–496.
- [67] Z. He, Y. Liang, L. Chen, I. Ahmad, and D. Wu, “Power-rate-distortion analysis for wireless video communication under energy constraints,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 5, pp. 645–658, May 2005.
- [68] C. Li, D. Wu, and H. Xiong, “Delay—Power-rate-distortion model for wireless video communication under delay and energy constraints,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 7, pp. 1170–1183, Jul. 2014.
- [69] R. Aparicio-Pardo and L. Sassatelli, “A green video control plane with fixed-mobile convergence and Cloud-RAN,” in *Proc. IEEE Int. Teletraffic Congr.*, 2017, pp. 28–36.
- [70] X. Corbillon, R. Aparicio-Pardo, N. Kuhn, G. Texier, and G. Simon, “Cross-layer scheduler for video streaming over MPTCP,” in *Proc. ACM 7th Int. Conf. Multimedia Syst.*, 2016, Art. no. 7.
- [71] B. Han, F. Qian, L. Ji, and V. Gopalakrishnan, “MP-DASH: Adaptive video streaming over preference-aware multipath,” in *Proc. ACM 12th Int. Conf. Emerg. Netw. Exp. Technol.*, 2016, pp. 129–143.
- [72] S. Chen, Z. Yuan, and G. M. Muntean, “An energy-aware multipath-TCP-based content delivery scheme in heterogeneous wireless networks,” in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2013, pp. 1291–1296.



Jiyian Wu (M'14) received the bachelor's degree from North China Institute of Science and Technology, Langfang, China, in June 2008, the master's degree from China University of Mining and Technology, Beijing, China, in June 2011, and the Ph.D. degree in computer science and technology from the Beijing University of Posts and Telecommunications, Beijing, China, in June 2014.

He was a Software Developer (C++) with Sinosoft Technologies Company Ltd., Beijing, China, from October 2010 to March 2014, a Postdoctoral Research Fellow with the Singapore University of Technology and Design, Singapore, from March 2014 to January 2016, and a Senior Software Engineer with OmniVision Technologies Singapore from January 2016 to September 2017. He is currently working on applying machine/deep learning algorithms in an industrial vision inspection system and real-time sensory data analysis. His research interests include video communication, video coding, forward error correction, heterogeneous wireless networks, concurrent multipath transfer, and energy optimization. His programming experiences/interests are in the areas of client-server video communication, embedded multimedia systems, peer-to-peer data transfer, and back-end server programs.



Rui Tan (SM'18) received the B.S. and M.S. degrees from Shanghai Jiao Tong University, Shanghai, China, in 2004 and 2007, respectively, and the Ph.D. degree in computer science from City University of Hong Kong, Hong Kong, in 2010. He is currently an Assistant Professor with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. Previously, he was a Research Scientist (2012–2015) and a Senior Research Scientist (2015) with Advanced Digital Sciences Center, a Singapore-based research center of the University of Illinois at Urbana-Champaign (UIUC), a Principle Research Affiliate (2012–2015) with Coordinated Science Lab of UIUC, and a Postdoctoral Research Associate (2010–2012) with Michigan State University. His research interests include cyber-physical systems, sensor networks, and pervasive computing systems. Prof. Tan received the Best Paper Awards from IPSN'17, CPSR-SG'17, and Best Paper Runner-Ups from IEEE PerCom'13 and IPSN'14.



Ming Wang received the bachelor's degree from the North China Institute of Science and Technology, Langfang, China, in June 2008 and the master's degree from the University of Science and Technology Beijing, Beijing, China, in January 2011. He is currently working toward the Ph.D. degree in computer science and technology in the State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China. He was a Software Engineer with the China Academy of Space Technology from 2011 to 2014 and a visiting Ph.D. candidate in Singapore University of Technology and Design from March 2015 to October 2016. His research interests include multimedia networking and image processing.