Received: 28 April 2020 | Revised: 16 November 2020 | Accepted: 25 November 2020

DOI: 10.4218/etrij.2020-0184

ORIGINAL ARTICLE

ETRI Journal WILEY

# Impact of playout buffer dynamics on the QoE of wireless adaptive HTTP progressive video

**Guannan Xie**[1] | **Huifang Chen**[1,2,3] | **Fange Yu**[1] | **Lei Xie**[1,2]

[1]College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China

[2]Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking, Hangzhou, China

[3]Zhoushan Ocean Research Center, Zhoushan, China

**Correspondence**
Huifang Chen, College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China.
Email: chenhf@zju.edu.cn

**Funding Information**
This research was supported by the Ministry of Industry and Information Technology of China (No. 2012ZX03001035-004), and the Science and Technology Department of Zhejiang Province (No. 2016C31060, No. 2018R52046).

The quality of experience (QoE) of video streaming is degraded by playback interruptions, which can be mitigated by the playout buffers of end users. To analyze the impact of playout buffer dynamics on the QoE of wireless adaptive hypertext transfer protocol (HTTP) progressive video, we model the playout buffer as a G/D/1 queue with an arbitrary packet arrival rate and deterministic service time. Because all video packets within a block must be available in the playout buffer before that block is decoded, playback interruption can occur even when the playout buffer is non-empty. We analyze the queue length evolution of the playout buffer using diffusion approximation. Closed-form expressions for user-perceived video quality are derived in terms of the buffering delay, playback duration, and interruption probability for an infinite buffer size, the packet loss probability and re-buffering probability for a finite buffer size. Simulation results verify our theoretical analysis and reveal that the impact of playout buffer dynamics on QoE is content dependent, which can contribute to the design of QoE-driven wireless adaptive HTTP progressive video management.

**KEYWORDS**
diffusion approximation, HTTP progressive video, playback interruption, playout buffer, QoE

## 1 | INTRODUCTION

Innovations and developments in multimedia equipment have made it possible to display high-quality videos. According to the Cisco visual network index [1], mobile data traffic will increase by seven times from 2017 to 2022, and videos will account for approximately 80% of mobile data traffic by 2022. Additionally, as video compression and wireless network technologies continue to develop rapidly, end users will become increasingly accustomed to resource-demanding multimedia services with superior quality [2].

Based on the strict deadlines for presentation in video streaming, the limitations of various network conditions and transport protocols may lead to video jitter or interruptions. Therefore, based on the stringent bandwidth and delay requirements of video streaming services, providing high-quality video to end users remains a challenge.

As the study of video streaming has shifted from the network-centric solutions to the human-centric solutions, analyzing user-perceived experiences regarding video quality has become a hot research topic [3–4]. Quality of experience (QoE) is considered as a more important metric than quality of service because it considers user-perceived quality directly [5]. Therefore, it is crucial to identify and quantify the major factors influencing the QoE of video streaming.

Over the past decade, determining the QoE of video streaming based on the user datagram protocol (UDP) has attracted significant research attention, and many QoE improvement techniques have been proposed. Most of these techniques focus on optimizing technical video parameters (for example, codecs, frame rates, and resolution), reducing the video distortion caused by network imperfections (for example, packet loss, delay, and jitter) [6] and adapting the playback rate and strategy at the receiver [7].

Nowadays, a huge amount of media content is available through progressive download techniques based on the hypertext transfer protocol (HTTP) and transmission control protocol (TCP) [8]. A media player at a receiver is typically deployed with a playout buffer that stores downloaded video packets. Therefore, a certain number of video packets can be buffered before decoding and displaying them, which delays the start of video playback for a short duration to alleviate the variable delay in packet transmission. Moreover, packet retransmission can mitigate the video distortion caused by packet loss, particularly in wireless networks.

However, video playout is still affected by the limitations of varying network conditions and transport protocols, which can be reflected by the packet arrival process at the playout buffer. Technical video parameters also influence the behaviors of the video playout buffer, including the video packet departure process and video playback. Therefore, it is important to understand how playout buffer dynamics influence the QoE of adaptive HTTP progressive video. In this study, we focus on buffer management.

The impact of network dynamics on user-perceived video quality was analyzed in [9] by modeling the playout buffer as a G/G/1 queue and deriving closed-form expressions for video quality using diffusion approximation. An upper bound for the analysis of video interruption probability under the effects of finite media file sizes was presented in [10].

Generally, a media file is partitioned into blocks, each of which contains several packets. If some packets within a block are not available at the receiver before that block is decoded, then a playback interruption will occur [11]. However, in related works, it has been assumed that a playback interruption will only occur if the playout buffer is empty. The effects of block length on user-perceived video quality have not been considered. Based on these motivations, we investigate the impact of playout buffer dynamics on the QoE of wireless adaptive HTTP progressive video.

In this study, we first develop an analytical framework to examine the effects of playout buffer dynamics on user-perceived video quality. With arbitrary packet arrival rates and a deterministic packet service time, the playout buffer is modeled as a $G/D/1/\infty$ queue in the case of an infinite buffer size and a $G/D/1/N$ queue in the case of a finite buffer size. By using the diffusion approximation method, we analyze the queue length evolution of the playout buffer. Closed-form expressions for user-perceived video quality are also derived. The obtained analytical results are validated using simulations. The primary contributions of this paper can be summarized as follows.

- Modeling: The setting of network protocols is generally determined by the mean and variance of the video packet arrival rate at the playout buffer. It is assumed that the inter-arrival interval of packets follows a general distribution. Because all video packets within a block must be available before that block is decoded, we assume that the packet departure process from the playout buffer is deterministic. If the buffer size is sufficiently large to store a complete media file, we model the playout buffer as a $G/D/1/\infty$ queue. Otherwise, we model it as a $G/D/1/N$ queue. Moreover, we consider a case in which video playback interruption occurs even when the playout buffer is non-empty. The effects of both the video block length and finite media file size on QoE are also studied.
- Solution: By using the diffusion approximation method, we analyze the queue length evolution of the playout buffer. We derive closed-form expressions for user-perceived video quality for an infinite buffer size in terms of the buffering delay, playback duration, and playback interruption probability. We also derive closed-form expressions for a finite buffer size in terms of the packet loss probability and re-buffering probability. The user-perceived video quality is characterized by the playout buffer dynamics, specifically the mean and variance of the video packet arrival rate, video packet departure rate, playback threshold, and video block length. The impact of playout buffer dynamics on the QoE of wireless adaptive HTTP progressive video can be evaluated based on our analytical results.
- Understanding: Our study provides insights into the most important aspects of playout buffer dynamics influencing the QoE of wireless adaptive HTTP progressive video. This enhanced understanding contributes to the design of QoE-driven network control and QoE-aware playout buffer management strategies.

The remainder of this paper is organized as follows. We present an overview of related work in Section 2. Section 3 presents our system model of the wireless adaptive HTTP progressive video and queue model of the playout buffer at the receiver. We describe our analytical framework for the cases of infinite and finite playout buffer sizes in Section 4 and verify the obtained analytical results using simulations in Section 5. The conclusions of this study are summarized in Section 6.

## 2 | RELATED WORK

Over the past decade, video streaming in wireless networks under varying conditions has been widely studied, and various

UDP- and TCP-based video streaming schemes have been proposed [12]. Our work is related to TCP-based video streaming and focuses on video quality from the perspective of end users.

User-perceived video streaming quality can be evaluated using both subjective and objective methods. Subjective evaluation is performed by humans, with the results graded according to the mean opinion score [13]. Although subjective evaluation directly reflects QoE, it is impractical based on large time and resource costs. Therefore, many studies have focused on objective evaluation methods for video streaming QoE. Khan and others [14] concluded that the effects of application and network parameters on video quality distortion are content dependent. Ghadiyaram and others [15] presented a continuous-time video QoE predictor that accounts for the interactions between stalling events, video content, and the state of the client-side data buffer. Qian and others [16] stated that the main factors relevant to QoE are video coding profiles, transmission errors, and buffering occurrence. Duanmu and others [17] constructed a video streaming database to study human responses to the effects of video compression, initial buffering, and playback interruption. To maximize video streaming QoE, an estimation framework based on re-buffering states and content encoding conditions was proposed in [18]. A video QoE-monitoring prototype platform considering bit rates and re-buffering was presented in [19]. Eswara and others [20] proposed a learning-based continuous QoE evaluation framework that parameterizes QoE evaluation in the playback and re-buffering states. Similar to [20], Eswara and others [21] used a recurrent neural network to predict continuous QoE based on a complex dataset considering playback indicators and re-buffering frequency. A pure buffer-based HTTP adaptive video streaming scheme for optimizing QoE was proposed in [22]. By comparing the effects of initial delay on playback interruption, Hossfeld and others [23] concluded that playback interruption must be avoided, but start-up delay increases as a result of pre-fetching. In [24], a buffer size optimization problem was proposed to avoid packet dropping caused by buffer overflow. Based on these studies, we conclude that the playout buffer has a significant influence on QoE, but none of these studies have examined the effects of playout buffer dynamics on video streaming QoE in an analytical framework.

In [9], by modeling the playout buffer as a G/G/1 queue, the authors derived closed-form expressions of the user-perceived video quality with an asymptotically large file size. In [10], the authors extended the work in [9] and proposed an analytical framework for video interruption probability with finite media file sizes. Xu and others [25] analyzed starvation behavior in a Markovian queue with finite packet arrival. However, in this related work, it was assumed that the playback interruption does not occur if the playout buffer is non-empty. The effects of both video block length and finite media file size on the interruption probability have not been addressed. In this study, we model the playout buffer as a G/D/1 queue under the assumption of constant video block length to analyze starvation behaviors. In particular, we consider the effects of the video block length on user-perceived video quality. In our conference paper [26], we discussed the effects of playout buffer dynamics on the QoE of adaptive HTTP progressive video in the case of an infinite buffer size. Further analysis of the work discussed in [26] and the case of a finite buffer size will be presented in this paper.

Another type of research has focused on diffusion approximation and its application to queuing systems [27]. Based on the diffusion approximation method, transient solutions for the queue length with different queue models were presented in [28] and [29]. Our work addresses the queue length evolution of the playout buffer using diffusion approximation, where the buffer is modeled as a G/D/1 queue.

## 3 | SYSTEM MODEL

We first describe the system architecture for wireless adaptive HTTP progressive video. We then present queue models for the playout buffer for infinite and finite buffer sizes.

Figure 1 presents the process of adaptive HTTP progressive video in wireless networks. Raw video content is encoded and stored in the streaming server. According to requests from end users, pre-stored video content is retrieved and segmented into packets. By using the HTTP/TCP/IP protocol suite, video packets are transmitted to user terminals under variable network conditions. A playout buffer is implemented at the receiver to mitigate the negative influences of network protocols and varying transmission conditions on video quality. For media playback, video packets are discharged from the buffer to be decoded and displayed to the end user. Here, we analyze the impact of playout buffer dynamics on the QoE of wireless adaptive HTTP progressive video from the perspective of the end user.

Based on the limitations of transport protocols and varying network conditions, video packets arrive at the playout buffer with variable delay. We assume that the inter-arrival interval of video packets is generally distributed with a mean of $1 / \lambda$ and variance of $\upsilon_a$.

During video encoding, raw video content is partitioned into video blocks with $w$ encoded packets per block. For a given codec, $w$ is determined by a group of pictures (GOP) value, the video content, and packet size. During playback, a video block must wait in the playout buffer until all packets belonging to that block arrive. If the GOP structure and frame rate are predetermined, then the video packets within a block can be assumed to depart at a constant rate. Therefore, we model the inter-departure interval for video packets as a distribution with a constant mean of $1/\mu$, where $\mu = w / t_0$, $t_0$ is the duration of the GOP.

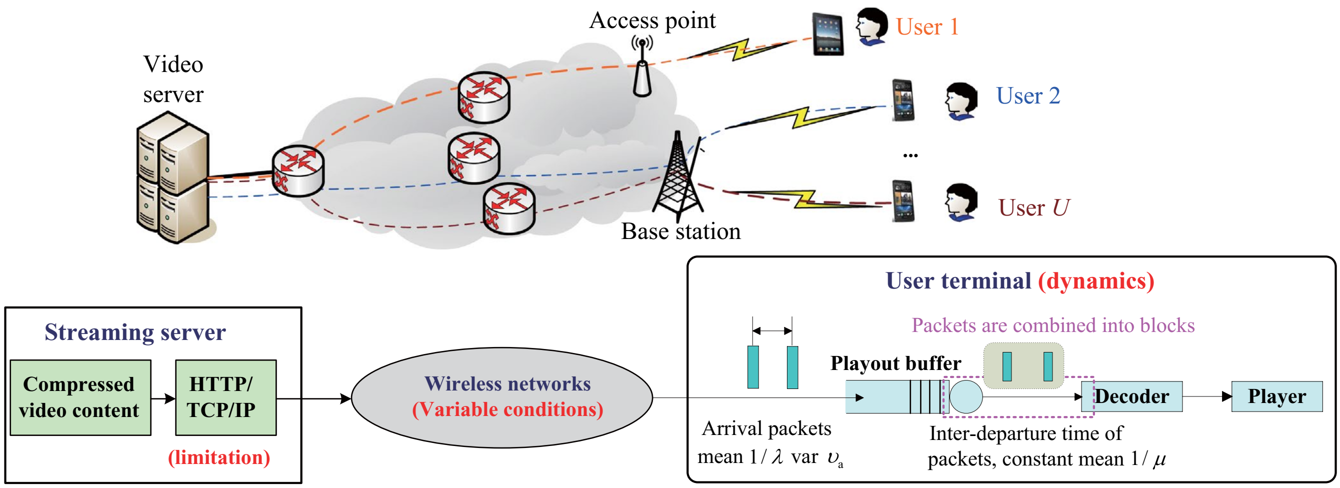



**FIGURE 1** Adaptive HTTP progressive video process in wireless networks

When the playout buffer is sufficiently large to store an entire media file, buffer evolution can be modeled as a G/D/1/∞ queue. Otherwise, it can be modeled as a G/D/1/$N$ queue.

To avoid video starvation, pre-fetching is adopted in most streaming applications [30]. In other words, some initial received video packets are stored, and media playback is delayed for a short duration (called the start-up delay). The number of pre-fetched packets is defined as the playback threshold. Because a video block with fewer than $w$ packets cannot be decoded, the playback threshold must be at least $w$. If the threshold is larger, the start-up delay will be longer, but the video starvation probability will also be lower. Therefore, to balance the trade-off between buffering delay and playback interruption, we set the playback threshold as $w + b$, where $b > 0$.

During video playback, to avoid playback interruption, video packets belonging to the next block to be decoded should be available in the playout buffer by the time the current block finishes playing. Therefore, in contrast with previous works, we consider a case in which playback interruption occurs, and the number of video packets left in the playout buffer is non-zero, but is less than $w$.

Because playout buffer occupancy is strongly related to start-up delay and video playback smoothness, we investigate HTTP progressive video QoE by analyzing the playout buffer evolution process for cases with infinite and finite buffer sizes when the playback statistics ($\lambda$, $\mu$, $\upsilon_a$), video block length $w$, and playback threshold $w + b$ are given.

# 4 | ANALYTICAL FRAMEWORK

In this section, we first discuss the playout buffer evolution process. We then develop an analytical framework to investigate user-perceived video quality in cases with infinite and finite buffer sizes.

## 4.1 | Case 1: Infinite buffer size

The buffer size can be assumed to be infinite when user terminals have large-volume storage and media files are much smaller than the playout buffer.

Figure 2 illustrates the playout buffer evolution process with an infinite buffer size during video playout. This process consists of two iterative phases: the buffering phase and playback phase.

In the buffering phase, video playback is frozen until the buffer is charged with $w + b$ (the playback threshold) packets. Then, the playback phase begins, and video packets are discharged from the buffer to be displayed. Under conditions with dynamic packet arrival, if the playback of the current video block ends and there are fewer than $w$ packets left in the playout buffer, then playback interruption occurs, and the buffering phase begins again. The buffering phase and playback phase iterate until the entire requested video file is downloaded or the user stops watching.

Let the random variable $D_m$ denote the duration of the buffering phase beginning with $m$ packets left in the buffer,

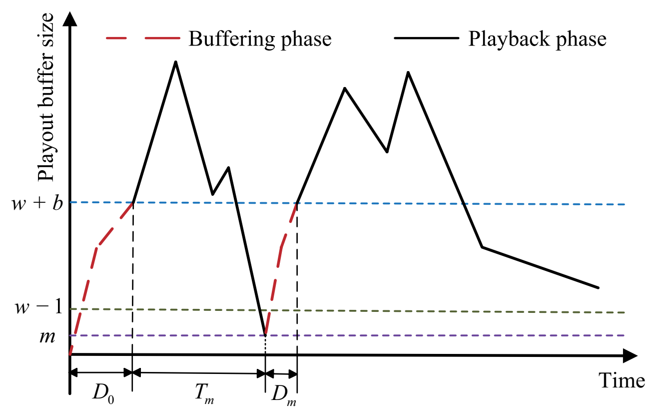



**FIGURE 2** Playout buffer evolution process with an infinite buffer size

where $m$ ranges from 0 to $w-1$. Because the packet departure of the playout buffer is deterministic, $m$ can be an arbitrary integer in $[0, w-1]$ with a probability $q_m = 1 / w$. Let the random variable $T_m$ denote the duration of the playback phase terminating with $m$ packets left in the buffer.

In this case, we evaluate the QoE of adaptive HTTP progressive video in terms of buffering delay and video playback smoothness. Buffering delay represents the duration that the user waits for playback to begin, which is denoted as $D_m$. Video playback smoothness is evaluated based on the playback duration $T_m$ and playback interruption probability $P_{\mathrm{I}}$.

As described previously, we model the infinite playout buffer as a G/D/1/∞ queue. Let $Y(t)$ denote the buffer length at time $t$. When using the diffusion approximation method, the discrete buffer length $Y(t)$ is replaced by a continuous diffusion process $X(t)$, which is modeled as the Brownia motion $dX(t) = X(T+dt) - X(t) = \beta dt + G(t)\sqrt{\alpha dt}$, where $G(t)$ is a white Gaussian process with zero mean and unit variance [28], $\beta$ and $\alpha$ are the instantaneous changes in the mean and variance of $X(t)$, respectively, which are defined as

$$\begin{cases} \beta = \lim\limits_{\Delta t \to 0} \dfrac{E\{X(t+\Delta t) - X(t)\}}{\Delta t} = \lambda - \mu, \\ \alpha = \lim\limits_{\Delta t \to 0} \dfrac{E\left\{[X(t+\Delta t) - X(t)]^2\right\} - [E\{X(t+\Delta t) - X(t)\}]^2}{\Delta t} = \lambda^3 v_a. \end{cases} \tag{1}$$

Let $p(x, t \mid x_0) = Pr\{x \le X(t) < x + dx \mid X(0) = x_0\}$ denote the conditional probability density function (PDF) of $X(t)$ at time $t$, where $x_0$ is the initial queue length. By using diffusion approximation, we characterize $p(x, t \mid x_0)$ as

$$\begin{cases} \dfrac{\partial p\left(x, t|x_0\right)}{\partial t} = \dfrac{\alpha}{2} \dfrac{\partial^2 p\left(x, t|x_0\right)}{\partial x^2} - \beta \dfrac{\partial p\left(x, t|x_0\right)}{\partial x}, \\ p\left(x, 0|x_0\right) = \delta\left(x - x_0\right), & t = 0, \\ p\left(x_e, t|x_0\right) = 0, & t > 0, \end{cases} \tag{2}$$

where $\delta(x)$ is the Dirac delta function. The third expression in (2) indicates that diffusion stops when $X(t) = x_e$.

### 4.1.1 | Buffering delay

The buffering phase begins with $m$ video packets and terminates with $w + b$ packets in the playout buffer. Therefore, the buffering delay $D_m = min\{t | X(0) = m, X(t) = w + b, t > 0\}$. Because the playback is frozen in the buffering phase ($\mu = 0$), we model the buffering phase as a diffusion process with $\beta_{\mathrm{D}} = \lambda$ and $\alpha_{\mathrm{D}} = \lambda^3 v_{\mathrm{a}}$. Therefore, the conditional PDF of $X(t)$ with $X(0) = m$ and $X(t) = w + b$, namely $p_{D_m}(x, t|m) = Pr\{x \le X(t) < x + dx \mid X(0) = m, X(\tau) < w + b,$ for $0 < \tau < t\}$, satisfies the diffusion equation as

$$\begin{cases} \dfrac{\partial p_{D_m}(x, t|m)}{\partial t} = \dfrac{\alpha_{\mathrm{D}}}{2} \dfrac{\partial^2 p_{D_m}(x, t|m)}{\partial x^2} - \beta_{\mathrm{D}} \dfrac{\partial p_{D_m}(x, t|m)}{\partial x}, \\ p_{D_m}(x, 0|m) = \delta(x - m), & t = 0, \\ p_{D_m}(w + b, t|m) = 0, & t > 0. \end{cases} \tag{3}$$

By solving (3), we obtain

$$p_{D_m}(x, t|m) = \frac{1}{\sqrt{2\pi\alpha_D t}} \left\{ \exp\left[ -\frac{(x - m - \beta_D t)^2}{2\alpha_D t} \right] - \exp\left\{ \frac{2\beta_D [(w+b) - m]}{\alpha_D} - \frac{[x - 2(w+b) + m - \beta_D t]^2}{2\alpha_D t} \right\} \right\}. \tag{4}$$

The conditional cumulative density function (CDF) of $X(t)$ can be calculated as

$$P_{D_m}(x, t|m) = \Phi\left( \frac{x - m - \beta_{\mathrm{D}} t}{\sqrt{\alpha_{\mathrm{D}} t}} \right) - \exp\left[ \frac{2\beta_{\mathrm{D}}(w + b - m)}{\alpha_{\mathrm{D}}} \right] \Phi\left( \frac{x - 2(w+b) + m - \beta_{\mathrm{D}} t}{\sqrt{\alpha_{\mathrm{D}} t}} \right), \tag{5}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy$.

The CDF and PDF of $D_m$ are

$$G_{D_m}(t) = 1 - \Phi\left( \frac{w + b - m - \beta_{\mathrm{D}} t}{\sqrt{\alpha_{\mathrm{D}} t}} \right) + \exp\left[ \frac{2\beta_{\mathrm{D}}(w + b - m)}{\alpha_{\mathrm{D}}} \right] \Phi\left( -\frac{w + b - m + \beta_{\mathrm{D}} t}{\sqrt{\alpha_{\mathrm{D}} t}} \right), \tag{6}$$

$$g_{D_m}(t) = \frac{w + b - m}{\sqrt{2\pi\alpha_{\mathrm{D}} t^3}} \exp\left\{ -\frac{\left[(w + b - m) - \beta_{\mathrm{D}} t\right]^2}{2\alpha_{\mathrm{D}} t} \right\}. \tag{7}$$

The moment-generating function (MGF) of $g_{D_m}(t)$ is

$$\phi_{D_m}(s) = \exp\left[\frac{w+b-m}{\alpha_D}\left(\beta_D - \sqrt{\beta_D^2 + 2s\alpha_D}\right)\right]. \tag{8}$$

Therefore, the mean and variance of $D_m$ can be calculated as

$$E\left(D_m\right) = -\frac{d}{ds}\phi_{D_m}(s)|_{s=0} = \frac{w+b-m}{\beta_D}, \tag{9}$$

$$Var\left(D_m\right) = \frac{d^2}{ds^2}\phi_{D_m}(s)|_{s=0} - E^2\left(D_m\right) = \frac{\alpha_D\left(w+b-m\right)}{\beta_D^3}. \tag{10}$$

The start-up delay is the duration of the first buffering phase with $m = 0$ in $D_m$. We also derive the average buffering delay of the buffering phase and its variance as

$$E(D) = \sum_{m=0}^{w-1} q_m E\left(D_m\right) = \sum_{m=0}^{w-1} \frac{1}{w}\cdot\frac{w+b-m}{\lambda} = \frac{w+2b+1}{2\lambda}, \tag{11}$$

$$Var(D) = E\left[\left(E\left[D_m\right] - E[D]\right)^2\right] = \frac{w^2-1}{12\lambda^2}. \tag{12}$$

#### 4.1.2 | Playback duration

The playback phase begins when the playout buffer is charged with $w+b$ video packets and ends with $m$ packets remaining in the playout buffer, as illustrated in Figure 2. Therefore, the playback duration is $T_m = min\ \{t|X(0) = w+b, X(t) = m, t > 0\}$. Because we model the playback phase as a diffusion process with $\beta_T = \lambda - \mu$ and $\alpha_T = \lambda^3 v_a$, the conditional PDF of $X(t)$ with $X(0) = w + b$ and $X(t) = m$, namely $p_{T_m}(x, t|w+b) = Pr\{x \le X(t) < x + dx | X(0) = w+b, X(\tau) > m,$ for $0 < \tau < t\}$, satisfies the diffusion equation as

$$\begin{cases} \dfrac{\partial p_{T_m}(x,t|w+b)}{\partial t} = \dfrac{\alpha_T}{2}\dfrac{\partial^2 p_{T_m}(x,t|w+b)}{\partial x^2} - \beta_T\dfrac{\partial p_{T_m}(x,t|w+b)}{\partial x}, \\ p_{T_m}(x,0|w+b) = \delta\left[x-(w+b)\right], \quad t=0, \\ p_{T_m}(m,t|w+b) = 0, \quad t>0. \end{cases} \tag{13}$$

By solving (13), we obtain

$$p_{T_m}(x,t|w+b) = \frac{1}{\sqrt{2\pi\alpha_T t}}\left\{\exp\left\{-\frac{\left\{[x-(w+b)]-\beta_T t\right\}^2}{2\alpha_T t}\right\} - \exp\left\{\frac{2\beta_T\left[m-(w+b)\right]}{\alpha_T} - \frac{\left[x-2m+(w+b)-\beta_T t\right]^2}{2\alpha_T t}\right\}\right\}. \tag{14}$$

Therefore, the conditional PDF and MGF of $T_m$ are

$$g_{T_m}(t) = -\frac{d}{dt}\int_m^\infty p_{T_m}(x,t|w+b)\,dx = \frac{w+b-m}{\sqrt{2\pi\alpha_T t^3}}\exp\left\{-\frac{\left[m-(w+b)-\beta_T t\right]^2}{2\alpha_T t}\right\}, \tag{15}$$

The mean and variance of $T_m$ are

$$\phi_{T_m}(s) = \int_0^\infty g_{T_m}(t)\,e^{-st}dt = \exp\left[\frac{m-(w+b)}{\alpha_T}\left(\beta_T + \sqrt{\beta_T^2 + 2s\alpha_T}\right)\right]. \tag{16}$$

$$E\left(T_m\right) = -\frac{d}{ds}\phi_{T_m}(s)|_{s=0} = \frac{w+b-m}{|\beta_T|},\ \beta_T < 0, \tag{17}$$

$$Var\left(T_m\right) = \frac{d^2}{ds^2}\phi_{T_m}(s)|_{s=0} - E^2\left(T_m\right) = \frac{\alpha_T\left(w+b-m\right)}{|\beta_T|^3},\ \beta_T < 0. \tag{18}$$

Therefore, the average playback duration of the playback phase and its variance are

$$E(T) = \sum_{m=0}^{w-1} q_m E\left(T_m\right) = \sum_{m=0}^{w-1}\frac{1}{w}\cdot\frac{w+b-m}{|\lambda-\mu|} = \frac{w+2b+1}{2|\lambda-\mu|},\ \lambda-\mu<0, \tag{19}$$

$$Var(T) = E\left[\left(E\left[T_m\right] - E[T]\right)^2\right] = \frac{w^2-1}{12(\lambda-\mu)^2},\ \lambda-\mu<0. \tag{20}$$

#### 4.1.3 | Playback interruption probability

Playback interruption occurs when the playback phase ends. Let $S$ be the media file size. The playback interruption probability with $X(0) = w + b$ and $X(t) = m$, namely $P_{I,m}$, can be calculated as

$$P_{I,m} = Pr\left\{t < S | X\ 0 = w+b, X\ t = m\right\} = \int_0^S g_{T_m}\ t\ dt$$
$$= \begin{cases} 1, & S \ge S_0, \beta_T \le 0, \\ \exp\left[-\dfrac{2\beta_T}{\alpha_T}\ w+b-m\right], & S \ge S_0, \beta_T > 0, \\ \exp\left[-\dfrac{2S\beta_T\ w+b-m\ +\ w+b-m^2}{2\alpha_T S}\right], & S < S_0, \beta_T \in \mathbb{R}, \end{cases} \tag{21}$$

where $S_0 = \frac{w+b-m}{2|\beta_T|}$ for $\beta_T \neq 0$ and $S_0 \to \infty$ for $\beta_T = 0$.

Because most media files satisfy $S \ge \frac{w+b}{2|\beta_T|}$, the playback interruption probability $P_I = E\left(P_{I,m}\right)$ can be calculated as

$$P_{\mathrm{I}}=\begin{cases}1, & \beta_{\mathrm{T}}=\lambda-\mu\le 0,\\ \dfrac{e^{-\frac{2\beta_{\mathrm{T}}}{\alpha_{\mathrm{T}}}b+1}\;1-e^{-\frac{2\beta_{\mathrm{T}}}{\alpha_{\mathrm{T}}}w}}{w\;1-e^{-\frac{2\beta_{\mathrm{T}}}{\alpha_{\mathrm{T}}}}}, & \beta_{\mathrm{T}}=\lambda-\mu>0.\end{cases} \tag{22}$$

According to (22), we find that the playback interruption must occur if the mean arrival rate $\lambda$ is less than the video playback rate $\mu$. Furthermore, even if $\lambda$ is greater than $\mu$, it is possible that video playback will be interrupted because of the dynamics of packet arrival and characteristics of video blocks.

## 4.2 | Case 2: Finite buffer size

Here, we discuss the case of a finite playout buffer size, where the media file is larger than the playout buffer. Figure 3 illustrates the playout buffer evolution process for a buffer size $N$ during video playout.

Because the start and stop conditions for the buffering phase are the same in both the cases of infinite and finite buffer sizes, the buffering delay analysis in Section 4.1.1 is also applicable in this case.

As illustrated in Figure 3, in the playback phase, the upper bound on the queue length is the playout buffer size $N$ and $N > w + b$. Once the playout buffer is full, the following video packets are dropped, which degrades the user-perceived video quality. Therefore, the packet loss probability caused by buffer overflow is an important performance metric in this case. $P_N$ denotes the packet loss probability, which is defined as

$$P_N=\lim_{t\to\infty} Pr\,\{X(t)=N\} \tag{23}$$

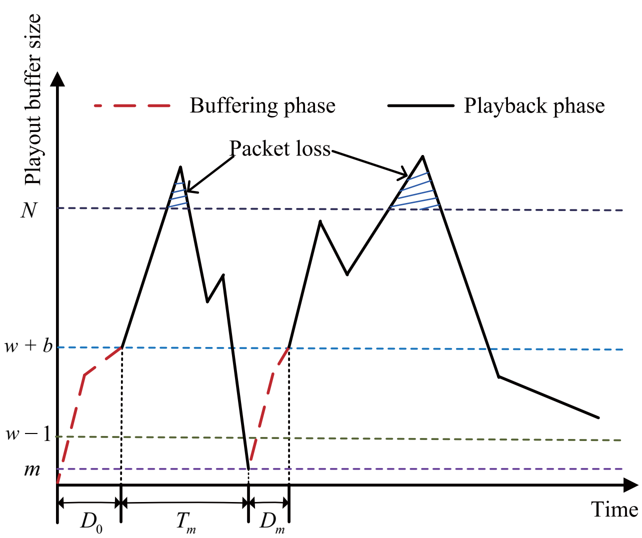



**FIGURE 3** Playout buffer evolution process with a buffer size $N$

Let $P_{\mathrm{BUF}}$ be the re-buffering probability, which is the probability that playback is interrupted and the playout buffer will enter the buffering phase immediately.

By using the diffusion approximation method, we model the playback phase in the case of a finite buffer size as

$$\begin{cases}\dfrac{\partial p(x,t|w+b)}{\partial t}=\dfrac{\alpha_{\mathrm{T}}}{2}\dfrac{\partial^2 p(x,t|w+b)}{\partial x^2}-\beta_{\mathrm{T}}\dfrac{\partial p(x,t|w+b)}{\partial x}\\ +\dfrac{1}{E(D)}P_{\mathrm{BUF}}\delta[x-(w+b)]+\mu P_N\delta(x-N+1),\\ \lim\limits_{x\to m}\left[\dfrac{\alpha_{\mathrm{T}}}{2}\dfrac{\partial p(x,t|w+b)}{\partial x}-\beta_{\mathrm{T}}p(x,t|w+b)\right]=\dfrac{1}{E(D)}P_{\mathrm{BUF}},\\ \lim\limits_{x\to N}\left[\dfrac{\alpha_{\mathrm{T}}}{2}\dfrac{\partial p(x,t|w+b)}{\partial x}-\beta_{\mathrm{T}}p(x,t|w+b)\right]=-\mu P_N,\end{cases} \tag{24}$$

where $\beta_{\mathrm{T}}=\lambda-\mu, \alpha_{\mathrm{T}}=\lambda^3 \upsilon_{\mathrm{a}}$, $m\in[0,w-1]$ with probability 1 / $w$, and $E(D)=\frac{w+2b+1}{2\lambda}$ according to (11).

The diffusion equation in (24) consists of two parts: the conditional PDF of $X(t)$ and the probabilities $P_N$ and $P_{\mathrm{BUF}}$ for the two boundaries when the buffer is full and in the buffering phase, respectively. $\frac{1}{E(D)}P_{\mathrm{BUF}}\delta[x-(w+b)]$ in (24) represents the probability that the buffer changes from buffering to playback, where 1 / $E(D)$ is the mean change rate. $\mu P_N\delta(x-N+1)$ represents the probability that the buffer queue length changes from $N$ to $N$—1. When the queue length is $N$, video packets depart at a rate $\mu$ and newly arriving video packets are dropped.

By solving (24) in the steady state, denoted $\lim\limits_{t\to\infty}\frac{\partial p(x,t|w+b)}{\partial t}=0$, we obtain

$$p(x,\infty|w+b)=\begin{cases}\dfrac{P_{BUF}}{\beta_T E(D)}(e^{rx}-1), & 0<x\le w+b,\\ \dfrac{P_{BUF}}{\beta_T E(D)}\left[1-e^{-r(w+b)}\right]e^{rx}, & w+b<x\le N-1,\\ \dfrac{\mu P_N}{\beta_T}\left[1-e^{r(x-N)}\right], & N-1<x\le N,\end{cases} \tag{25}$$

$$P_N=\frac{e^{r(N-1)}\left[1-e^{-r(w+b)}\right]}{\mu E(D)(1-e^{-r})}P_{\mathrm{BUF}}, \tag{26}$$

where $r=\frac{2\beta_{\mathrm{T}}}{\alpha_{\mathrm{T}}}=\frac{2(\lambda-\mu)}{\lambda^3\upsilon_{\mathrm{a}}}$.

Because $\int_0^N p(x,t|w+b)\,dx+P_{\mathrm{BUF}}+P_N=1$, we can obtain the packet loss probability $P_N$ and re-buffering probability $P_{\mathrm{BUF}}$ as follows:

$$P_N = \left\{ \frac{-\mu^2 (w+b)(1-e^{-r})}{\lambda(\lambda-\mu)e^{r(N-1)}\left[1-e^{-r(w+b)}\right]} + \frac{\lambda}{\lambda-\mu} \right\}^{-1}, \tag{27}$$

$$P_{\mathrm{BUF}} = \left\{ \frac{2\lambda^2 e^{r(N-1)}\left[1-e^{-r(w+b)}\right]}{\mu(\lambda-\mu)(w+2b+1)(1-e^{-r})} - \frac{2\mu(w+b)}{(\lambda-\mu)(w+2b+1)} \right\}^{-1}. \tag{28}$$

## 5 | NUMERICAL RESULTS AND DISCUSSION

We validate the analytical results discussed above using MATLAB-based simulations in this section. In the following figures, "analysis" and "simulation" are abbreviated as "ana" and "sim," respectively.

### 5.1 | Simulation setup

We chose two video clips, namely "Star War IV" and "Silence of the Lambs," from [31]. Both clips are in MPEG-4 format with the same frame rate of 25 fps. The video duration is 1 h, and the resolution follows the quarter common intermediate format. The GOP structure is IBBPBBPBBPBB. The duration of the GOP, denoted by $t_0$, is 0.48 s. The video frame statistics are listed in Table 1, which shows that the frame size is related to the quality of the videos. The video frames are segmented into IP packets with a data payload size of 1200 bytes. $\lambda / \mu$ is defined as the traffic density.

### 5.2 | Case 1: Infinite buffer size

In this subsection, we present the analytical and simulation results for the start-up delay and video playback interruption probability for an infinite buffer size. The video packet arrival process is assumed to follow a Poisson distribution and the inter-arrival interval of the video packets is exponentially distributed with a mean of $1/\lambda$ and variance of $1/\lambda^2$.

Figure 4 shows the CDFs of the start-up delays for different traffic densities, where $b = 3w$. From Figure 4, we can observe that as the traffic density increases, the mean start-up delay decreases because the corresponding CDF curve moves to the left and the variance decreases because the CDF curve becomes compressed in width. Moreover, for a given traffic density, as the video quality enhances, the mean start-up delay increases because the corresponding CDF curve moves to the right and the variance decreases because the CDF curve becomes compressed in width.

**TABLE 1** Video frame statistics

| Video clip | Quality | Mean frame size (bytes) | $\mu = w/t0$ (packets per second) |
|---|---|---|---|
| Star Wars IV | Low | 270 | 6.25 |
| Silence of the Lambs | high | 2900 | 60.417 |

In addition, to limit the start-up delay to less than 2 s, the mean packet arrival rate $\lambda$ must be at least 2 times and 1.2 times the playback rate $\mu$ for low- and high-quality videos, respectively. Therefore, for the given playback threshold, more traffic density should be allocated to the low-quality video than the high-quality video.

The CDFs of the start-up delays for different playback thresholds are presented in Figure 5, where the traffic density, $\lambda / \mu$, is set to 2 and 1.2 for the low- and high-quality videos, respectively.

From Figure 5, we can observe that for a given traffic density with different video qualities, both the mean and

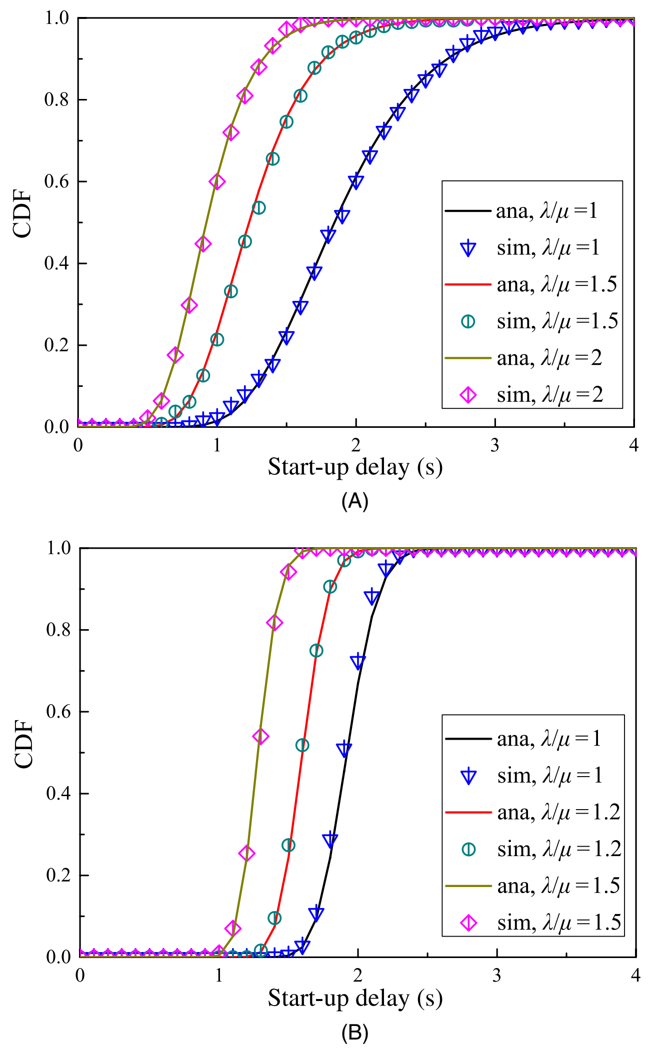



**FIGURE 4** CDFs of the start-up delay for different traffic density: (A) low-quality video and (B) high-quality video

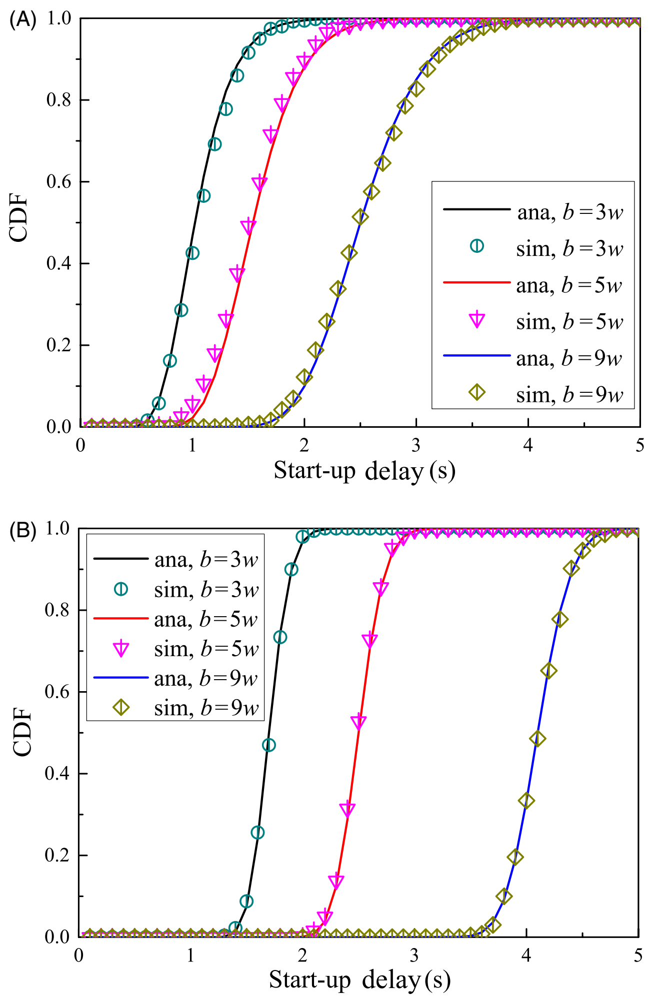



**FIGURE 5** CDFs of the start-up delays for different playback thresholds: (A) low-quality video with $\lambda/\mu = 2$ and (B) high-quality video with $\lambda/\mu = 1.2$

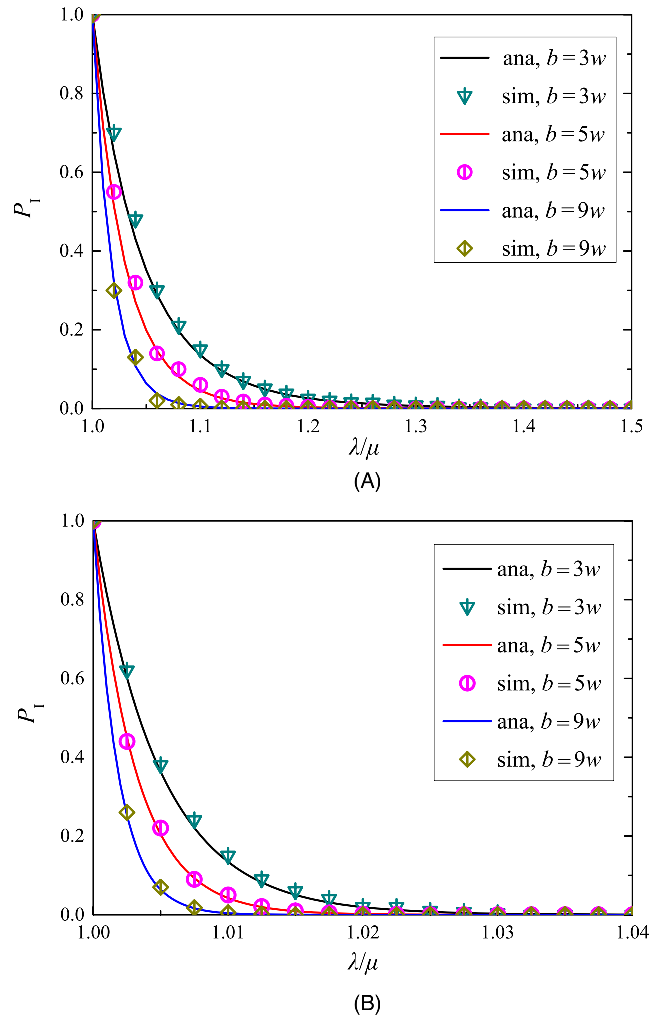



**FIGURE 6** Interruption probabilities for different playback thresholds and traffic densities: (A) low-quality video and (B) high-quality video

variance of the start-up delay increase as the playback threshold increases because both the corresponding CDF curves in Figure 5A and 5B move to the right and expand in width. Therefore, for a given traffic density, a small playback threshold is preferable to achieve a small mean and variance for the start-up delay.

Figure 6 presents the playback interruption probabilities for different playback thresholds and traffic densities. One can see that the playback interruption probability decreases as the traffic density increases. In addition, the playback interruption probability of the playout buffer with a larger playback threshold decreases more rapidly as the traffic density increases. Therefore, for a given traffic density with different video qualities, a larger playback threshold is preferable to achieve a lower playback interruption probability. Additionally, for a given playback threshold, to obtain a consistent playback interruption probability, the traffic density for the low-quality video should be greater than that for the high-quality video.

Therefore, according to Figures 5 and 6, the playback threshold of the playout buffer must be set properly to achieve a proper trade-off between the start-up delay and video playback interruption probability. Figures 4 and 6 demonstrate that the impact of playout buffer dynamics on QoE is content dependent and that more traffic density should be allocated to low-quality videos than high-quality videos.

Moreover, the obtained analytical results agree well with all of the simulation results, thereby verifying the accuracy of our theoretical analysis for an infinite buffer size.

## 5.3 | Case 2: Finite buffer size

In this subsection, we validate the analytical and simulation results for the packet loss probability and re-buffering probability in the case of a finite buffer size. The arrival process of video packets is still assumed to follow a Poisson distribution.

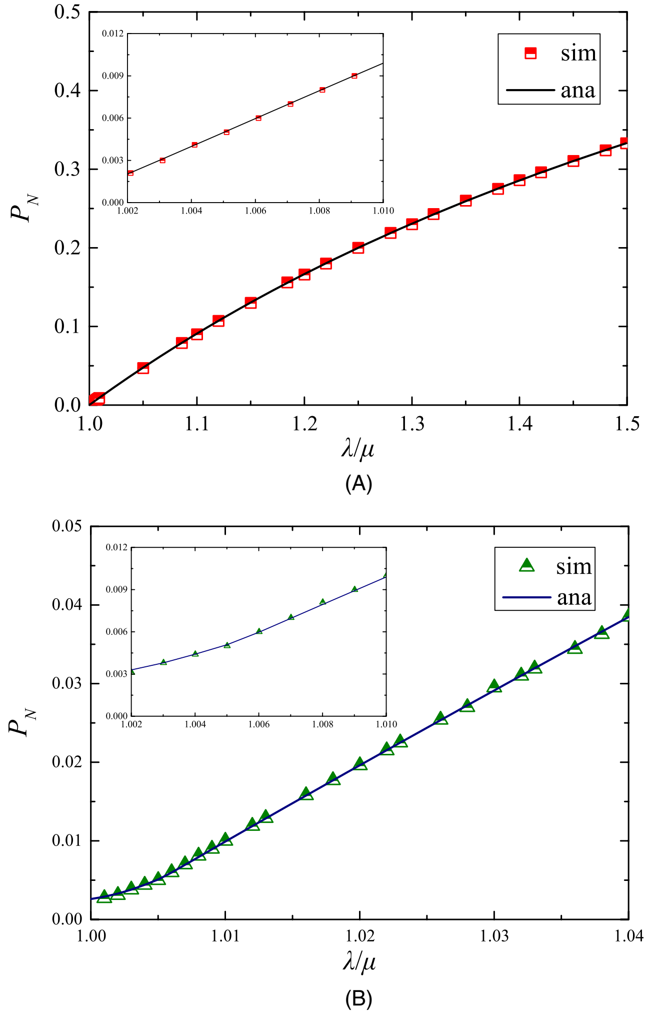



**FIGURE 7** Packet loss probabilities for different traffic densities: (A) low-quality video and (B) high-quality video

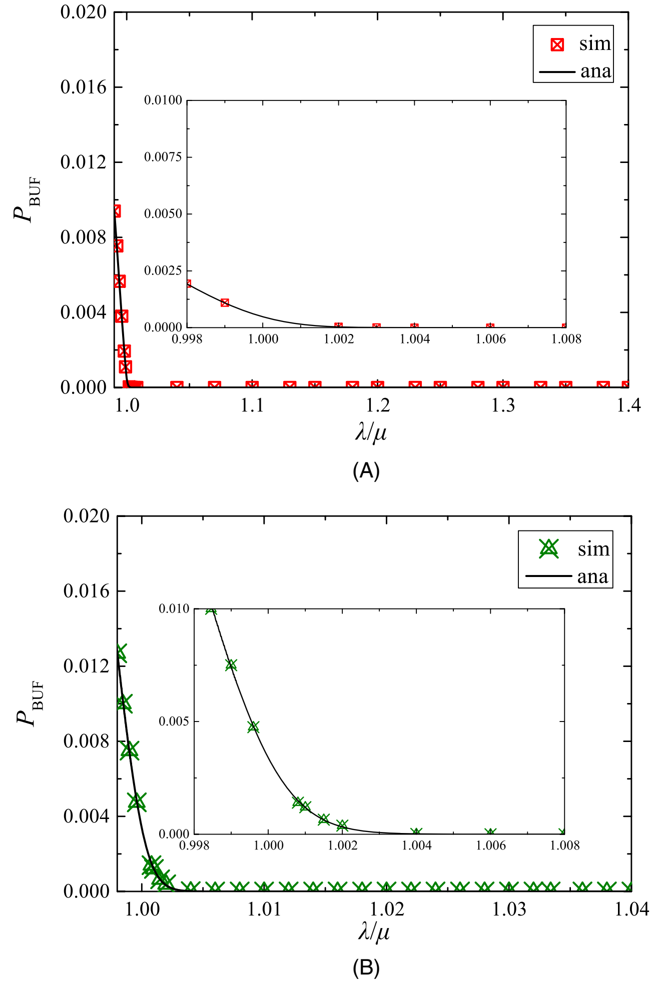



**FIGURE 8** Re-buffering probabilities for different traffic densities: (A) low-quality video and (B) high-quality video

Figure 7 presents the impact of traffic density on the packet loss probability, $P_N$, where $N = 1000$ packets and $b = 5w$. In Figure 7, one can see that the packet loss probability increases as the traffic density increases. With a higher traffic density, more video packets are stored in the playout buffer, and the buffer overflows more easily.

In addition, given a traffic density near one, which indicates that the packet arrival rate is almost the same as the packet departure rate, the packet loss probability for the low-quality video is smaller than that for the high-quality video.

The effects of traffic density on the re-buffering probability, $P_{BUF}$, are presented in Figure 8, where $N = 1000$ packets and $b = 5w$. From Figure 8, we can observe that the re-buffering probability decreases as the traffic density increases. With a higher traffic density, more video packets are stored in the playout buffer, and the video playback becomes more difficult to interrupt.

Moreover, for different video qualities, the effect of traffic density on the re-buffering probability for the low-quality video is weaker than that for the high-quality video. In addition, for a given traffic density, the re-buffering probability for the low-quality video is smaller than that for the high-quality video.

Figures 7 and 8 show that for a given playback threshold and buffer size, the playout buffer dynamics must be controlled properly to achieve a suitable trade-off between packet loss probability and re-buffering probability.

Moreover, it is clear that the analytical results agree well with the simulation results in the case of a finite buffer size.

## 6 | CONCLUSIONS

In this study, we analyzed the influence of playout buffer dynamics on the QoE of wireless adaptive HTTP progressive video. We modeled the playout buffer at the receiver as a G/D/1/∞ queue for an infinite buffer size and G/D/1/$N$ queue for a finite buffer size, with arbitrary packet arrival and deterministic service time. By diffusion approximation, we developed an analytical framework to derive closed-form expressions for user-perceived video quality. Simulation results verified the accuracy of our theoretical analysis.

This work revealed that the impact of playout buffer dynamics on QoE is content dependent and that the effect on high-quality videos is stronger than that on low-quality videos. We also determined that well-controlled playout buffer dynamics contribute to achieving high user-perceived video quality.

In the future, we will study the influence of playout buffer dynamics on the QoE of wireless adaptive HTTP progressive video streaming with a scalable bit rate and design QoE-driven wireless adaptive HTTP progressive video management schemes based on the proposed analytical framework.

## ORCID

*Guannan Xie* https://orcid.org/0000-0002-4003-3603
*Huifang Chen* https://orcid.org/0000-0002-1366-1030

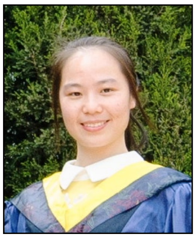
## REFERENCES

1. Cisco, *Cisco visual networking index: Global mobile data traffic forecast update, 2017-2022 White Paper*, USA, 2019.
2. A. A. Barakabitze et al., *QoE management of multimedia streaming services in future networks: A tutorial and survey*, IEEE Commun. Surv. Tutor. **22** (2020), 526–565.
3. F. Dobrian et al., *Understanding the impact of video quality on user engagement*, SIGCOMM Comput. Commun. Rev. **41** (2011), 362–373.
4. U. Reiter, *Overall perceived audiovisual quality−What people pay attention to*, in Proc. IEEE Int. Symp. Consum. Electron. (Singapore, Singapore), June 2011, pp. 513–517.
5. L. A. Rowe and R. Jain, *ACM SIGMM retreat report on future directions in multimedia research*, ACM Trans. Multimedia Comput. Commun. Appl. **1** (2005), 3–13.
6. K. Piamrat, et al., *Quality of experience measurements for video streaming over wireless networks*, in Proc. Int. Conf. Inform. Technol. New Generations (Las Vegas, NA, USA), Apr. 2009, pp. 1184–1189.
7. J. Yang et al., *Online buffer fullness estimation aided adaptive media playout for video streaming*, IEEE Trans. Multimedia **13** (2011), 1141–1153.
8. T. D. Pessemier et al., *Quantifying the influence of rebuffering interruptions on the user's quality of experience during mobile video watching*, IEEE Trans. Broadcast. **59** (2013), 47–61.
9. T. H. Luan et al., *Impact of network dynamics on user's video quality: Analytical framework and QoS provision*, IEEE Trans. Multimedia **12** (2010), 64–78.
10. A. Anttonen and A. Mämmelä, *Interruption probability of wireless video streaming with limited video lengths*, IEEE Trans. Multimedia **16** (2014), 1176–1180.
11. A. ParandehGheibi et al., *Avoiding interruptions−A QoE reliability function for streaming media applications*, IEEE J. Sel. Areas Commun. **29** (2011), 1064–1074.
12. W. Pu et al., *Dynamic adaptive streaming over HTTP from multiple content distribution servers*, in Proc. IEEE Global Telecommun. Conf. (Houston, TX, USA), Dec. 2011, pp. 1–5.
13. K. Seshadrinathan et al., *Study of subjective and objective quality assessment of video*, IEEE Trans. Image Process. **19** (2010), 1427–1441.
14. A. Khan et al., *QoE prediction model and its application in video quality adaptation over UMTS networks*, IEEE Trans. Multimedia **14** (2012), 431–442.
15. D. Ghadiyaram et al., *Learning a continuous-time streaming video QoE model*, IEEE Trans. Image Process. **27** (2018), 2257–2271.
16. L. Qian, et al., *SVM-based QoE estimation model for video streaming service over wireless networks*, in Proc. Int. Conf. Wirel Commun. Signal Process. (Nanjing, China), Oct. 2015, pp. 1–6.
17. Z. Duanmu et al., *A quality-of-experience index for streaming video*, IEEE J. Sel. Topics Signal Process. **11** (2017), 154–166.
18. T. Kimura et al., *QUVE: QoE maximizing framework for video-streaming*, IEEE J. Sel. Topics Signal Process. **11** (2017), 138–153.
19. T. Mangla et al., *VideoNOC: Assessing video QoE for network operators using passive measurements*, in Proc. ACM Multimedia Syst. Conf. (Amsterdam, Netherlands), June 2018, pp. 101–112.
20. N. Eswara et al., *A continuous QoE evaluation framework for video streaming over HTTP*, IEEE Trans. Circuits Syst. Video Technol. **28** (2018), 3236–3250.
21. N. Eswara et al., *Streaming video QoE modeling and prediction: A long short-term memory approach*, IEEE Trans. Circuits Syst. Video Technol. **30** (2020), 661–673.
22. W. Huang et al., *Buffer state is enough: Simplifying the design of QoE-aware HTTP adaptive video streaming*, IEEE Trans. Broadcast. **64** (2018), 590–601.
23. T. Hossfeld et al., *Initial delay vs. interruptions: Between the devil and the deep blue sea*, in Proc. Int. Workshop Qual. Multimedia Experience (Yarra Valley, Australia), July 2012, pp. 1–6.
24. H. Bobarshad et al., *Analytical modeling for delay-sensitive video over WLAN*, IEEE Trans. Multimedia **14** (2012), 401–414.
25. Y. Xu et al., *Analysis of buffer starvation with application to objective QoE optimization of streaming services*, IEEE Trans. Multimedia **16** (2014), 813–827.
26. F. Yu et al., *Impact of end-user playout buffer dynamics on HTTP progressive video QoE in wireless networks*, in Proc. IEEE Annual Int. Symp. Personal Indoor Mobile Radio Commun. (Washington, DC, USA), Sept. 2014, pp. 1996–2001.
27. A. Duda, *Transient diffusion approximation for some queueing systems*, in Proc. ACM SIGMETRICS Conf. Measurement Modeling Computer Syst. (Minneapolis, MN, USA), Aug. 1983, pp. 118–128.
28. H. Kobayashi, *Application of the diffusion approximation to queuing networks: Part I equilibrium queue distributions*, J. ACM **21** (1974), 316–328.
29. E. Gelenbe, *On approximate computer system models*, J. ACM **22** (1975), 261–269.
30. W. Zhang et al., *QoE-driven cache management for HTTP adaptive bit rate streaming over wireless networks*, IEEE Tran. Multimedia **15** (2013), 1431–1445.
31. F. H. P. Fitzek and M. Reisslein, *MPEG-4 and H.263 video traces for network performance evaluation*, IEEE Netw. **15** (2001), 40–54.

## AUTHOR BIOGRAPHIES

**Guannan Xie** received his BS degree in communication engineering from Zhejiang University, Hangzhou, China, in 2012. He is currently a PhD candidate majoring in communications and information systems at Zhejiang University. His current research interests include heterogeneous wireless networks, network selection, concurrent multipath transmission, and multimedia streaming.

**Huifang Chen** received her BS degree in electronic engineering, MS degree in communications and electronic systems, and PhD in communications and information systems from Zhejiang University, Hangzhou, China, in 1994, 1997, and 2000, respectively. Since 2000, she has been working at Zhejiang University, where she is currently a full professor in the College of Information Science and Electronic Engineering. Her current research interests include wireless networks, underwater acoustic networks, adaptive networks, and network security. She is a member of IEEE and ACM, and a senior member of the China Institute of Communications.



**Fange Yu** received her BS degree in communication engineering from Southeast University, Nanjing, China, in 2012, and her MS degree in information and communication engineering from Zhejiang University, Hangzhou, China, in 2015. She is currently a system engineer at China Mobile Inc., Hangzhou, China. Her research interests include wireless multimedia streaming and resource management in heterogeneous wireless networks.

**Lei Xie** received his BS degree in electronic engineering, MS degree in communications and electronic systems, and PhD degree in communications and information systems from Zhejiang University, Hangzhou, China, in 1994, 1997, and 2002, respectively. Since 1997, he has been with Zhejiang University, where he is currently an associate professor in the College of Information Science and Electronic Engineering. His current research interests include information theory and coding, network security, and multimedia streaming in heterogeneous wireless networks. He is a member of IEEE and a senior member of the China Institute of Communications.