Parametric Planning Model for Video Quality Evaluation of IPTV Services Combining Channel and Video Characteristics

Jiarun Song, Fuzheng Yang, Member, IEEE, Yicong Zhou, Senior Member, IEEE and Shan Gao

Abstract—Parametric planning models are designed for estimating the video quality, which can be applied to effective plan, implementation, and management of network video applications and communication networks. However, different from the bitstream-based evaluation models, the planning models are not allowed to exploit the video streams, with only limited information available for use, i.e., a few general parameters pre-determined by the service providers and network operators. In this paper, a parametric planning model combining channel and video characteristics is proposed to estimate the video distortion caused by packet loss for IPTV services. More specifically, the probability distribution of the channel states is determined by detailed analysis of the channel characteristics. Then, considering the influence of burst packet loss and the temporal dependence between frames, several sequence-level and frame-level parameters for video quality evaluation are derived from the perspective of the probability distribution of the channel states. Utilizing these parameters, the proposed model approximates the video quality considering the effects of direct packet loss and error propagation. Experimental results show that the proposed model has a superior performance for video quality estimation than three commonly used parametric planning models.

Index Terms—Video quality assessment, video streaming applications, network planning, QoE planning, planning model

I. INTRODUCTION

Recent years have witnessed an increasing proliferation of IPTV services. According to Cisco’s report, the global IP video traffic will be 82 percent of all consumer Internet traffic by 2020, up from 70 percent in 2015 [1]. In order to achieve a high level of user satisfaction for IPTV services, it is crucial to objectively predict the video quality for system design, network and Quality of Experience (QoE) planning.

This work was supported by the National Science Foundation of China under Grant 61601349, 61371089, 61571337, and 61601348, and by the 111 Project B08038.

Jiarun Song is with Collaborative Innovation Center of Information Sensing and Understanding and ISN, Xidian University, China (e-mail: jsong@xidian.edu.cn).

Fuzheng Yang is with the State Key Laboratory of ISN, Xidian University, China, and with the School of Electrical and Computer Engineering, Royal Melbourne Institute of Technology, Melbourne, VIC 3001, Australia (e-mail: fyang@mail.xidian.edu.cn).

Yicong Zhou is with Department of Computer and Information Science, University of Macau, Macau, China (e-mail: yicongzhou@umac.mo).

Shan Gao is with Media Technology Laboratory, Huawei Technologies Co., Ltd., Shenzhen, China (e-mail: simon.gaoshan@huawei.com).

In terms of the employed information, objective quality assessment models for network video can be classified into five categories: parametric planning models, packet-layer models, bitstream-layer models, media-layer models, and hybrid models [5], as illustrated in Fig.1. Different from the packet-layer model, the bitstream-layer model and the hybrid model which are usually used for quality monitoring [6], the parametric planning model is mainly applied to network and service planning. It is initially designed for service planners to identify beforehand how the video quality will be in a certain application and network parameters setting, to avoid over-engineering the applications, terminals, and networks while guaranteeing user’s satisfaction [7]. More specifically, the video quality is estimated using a priori of parameters, and then the appropriate service and network parameters are chosen and deployed in practice according to the video quality. Thus, the parametric planning model is important and helpful for multimedia service providers and network operators.

However, unlike other kinds of assessment models that evaluate the video quality by exploiting the information of bitstream or media signals, as shown in Fig. 1, the parametric planning model estimates the video quality without resorting to the actual video streams during the planning phase, and uses only a few empirical parameters (e.g., coding bitrate, packet loss rate, and so on) supplied by the service providers and network operators. In such a case, the parametric planning model cannot obtain the detailed information about video coding (e.g., frame type, quantization parameter, motion vector, etc.), video content (e.g., temporal complexity, spatial complexity, pixel values, etc.), as well as packet loss (e.g., the actual position of packet loss, the number of lost packet, etc.). How to accurately estimate the video quality using parametric planning model challenging and still remains as an open issue.

Targeting service planning, a parametric planning model was standardized by ITU-T Recommendation G.1070 for videophones, where the video quality was calculated by using application and network parameters, such as the bit rate, frame rate and packet loss rate [8]. Considering the influence of the burst packet loss on the video quality, as shown in Table I, the burst length [9] and the number of lost packets in a row [10] were employed to estimate the video quality, respectively. The model in [9] was an update version of the ITU-T G.1070 model. However, all these parametric planning models simply use a few statistical parameters to estimate the video quality and fail to carefully analyze the influence of the packet loss on the video quality benchmarking and monitoring [1]-[4].
coded frame and video streaming. In recent years, ITU-T Study Group 12 studied a new parametric planning model (G.OMVAS) for video streaming applications, which focuses on evaluating the impact of typical IP network impairments on the video quality [11]. The corresponding parametric planning model was presented in the ITU-T Recommendation G.1071 [12], whose formula and outputs are in accordance with those of packet layer model in ITU-T Recommendation P.1201 [13]. Particularly, the ITU-T G.1071 model provides a set of rules to convert the planning parameters into the forms of P.1201 inputs since these packet-layer inputs are not available in the planning phase. Though the influence of packet loss on the video streaming was studied in this model, the correlation between individual packet loss events is still not taken into account.

In practice, the packet loss process in the wired and wireless channels often exhibits finite temporal dependency and can be well characterized via a finite-state Markov model [14]-[17]. In this paper, a parametric planning model is proposed to evaluate the video quality of IPTV services, which matches the G.OMAVS framework outlined by ITU-T SG12. The model covers the H.264/AVC coded video transmitted over channels modelled by a four-state Markov chain. Unlike most traditional methods which directly map the statistical coding and network parameters (e.g., coded bit-rate, packet loss rate, etc.) to video quality, the proposed model evaluates video quality combining channel and video characteristics. Specifically, the probability distribution of the channel states is calculated analyzing channel characteristics, which can better clarify different packet loss behaviors. Due to the fact that the burst packet loss and the temporal dependence between frames will lead to nonrandom frame distortion, several sequence- and frame-level parameters for quality evaluation of IPTV services are derived from the perspective of probability distribution of the channel states. Utilizing these parameters, the proposed model can better indicate the effects of direct packet loss and error propagation on the video streaming. It is noted that for IPTV services, the video streams are more sensitive to distortion caused by packet loss, while one-way delay and jitter are generally not problematic since the Set-Top-Box (STB) is able to provide proper de-jitter buffers [4]. Thus, this research is focused on how to investigate the influence of packet loss on the video quality for IPTV services.

The remainder of this paper is organized as follows. Section II describes the framework of the proposed parametric planning model. In Section III, the impacts of packet loss on video quality are evaluated. Performance evaluation and conclusion are given in Section IV and V, respectively.

II. FRAMEWORK OF THE PROPOSED PARAMETRIC PLANNING MODEL

The framework of the proposed parametric planning model is illustrated in Fig. 2. The proposed model consists of five modules: input parameter module, coefficient database module, parameter analysis module, quality-evaluation module for coding distortion, and quality-evaluation module for distortion caused by packet loss.
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2638621, IEEE Transactions on Multimedia

average impaired ratio of each frame can be estimated module. For instance, considering the influence of the packet quality evaluation can be deduced by the parameter analysis channel models, some other useful information for video (3) evaluate the video quality.

database module will provide the corresponding coefficients to slicing, zero motion error concealment), the coefficient according to the input parameters (e.g., H.264 codec, 720P, coefficients of the parametric planning model. Therefore, different video codecs and resolutions may have different video resolutions or many others. The video streaming with which is designed to cope with the different video codec type, the input parameter module includes the input parameters of the planning model. Different from the bitstream-based or pixel-based video quality assessment model that use the information extracted from an available video streaming, the inputs of the planning model are the statistical parameters supplied by the network operators and service providers during the planning phase. There are only a few parameters that can be used in the planning model. According to the proposal of ITU-T SG12 for the parametric planning models [18], the input parameters can be divided into two categories: the input parameters for quality estimation and those for the coefficient database. More specifically, the input parameters for quality estimation are used to evaluate the video quality, while the input parameters for the coefficient database are used to determine which group of model coefficients should be chosen to evaluate the video quality. The detailed input parameters for the parametric planning models are listed in Table II.

TABLE II

<table>
<thead>
<tr>
<th>Category</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>INPUT PARAMETERS FOR THE PARAMETRIC PLANNING MODEL.</td>
<td></td>
</tr>
<tr>
<td>Bit-rate, frame rate, packet loss rate, averaged burst packet loss length, packet size, length of GOP, number of slice, and characteristics of channel model (state transition probability).</td>
<td></td>
</tr>
<tr>
<td>Codec type (H.264, MPEG4), video resolution, packet loss concealment (freezing, slicing)</td>
<td></td>
</tr>
</tbody>
</table>

(2) Coefficient database module
The coefficient database module stores a set of coefficients, which is designed to cope with the different video codec type, video resolutions or many others. The video streaming with different video codecs and resolutions may have different coefficients of the parametric planning model. Therefore, according to the input parameters (e.g., H.264 codec, 720P, slicing, zero motion error concealment), the coefficient database module will provide the corresponding coefficients to evaluate the video quality.

(3) Parameter analysis module
Combining input parameters with characteristics of the channel models, some other useful information for video quality evaluation can be deduced by the parameter analysis module. For instance, considering the influence of the packet loss on the video streaming, the average frame loss frequency, the expectation of the number of impaired frames and the average impaired ratio of each frame can be estimated combining the characteristics of the channel models. These parameters can be further used to estimate the video quality. The detailed procedure of the parameter analysis module is one of the contributions of the proposed planning model.

(4) Quality-evaluation module for coding distortion
This module is used to evaluate the video coding distortion. According to [19], [20], the video quality affected by the coding distortion is closely related to the coded bitrate and frame rate. It has been effectively evaluated as follows:

$$Q = \begin{cases} v_1 \left( 1 - \frac{1}{1 + (B_r / v_3)^v_2} \right), & F_r \geq 30 \\ v_1 \left( 1 - \frac{1}{1 + (B_r / v_3)^v_2} \right) \left( 1 - v_1 \ln \left( \frac{30}{F_r} \right) \right), & F_r < 30 \end{cases}$$

where $Q$ is the coding quality, $F_r$ is the video frame rate, $B_r$ is the average number of bits for coding a frame, which can be obtained by the ratio of the coded bit-rate and $F_r$. $v_1$, $v_2$, $v_3$ and $v_4$ are empirical parameters. Considering that all the mentioned parameters are available during the planning phase, the proposed parametric planning model utilizes equation (1) to estimate the video coding quality as well, where the parameters of $v_1$, $v_2$, $v_3$, $v_4$ and $v_5$ are retrained using the current database.

(5) Quality-evaluation module for distortion caused by packet loss
This module evaluates the video quality when the video transmission in the presence of channel errors. According to the study in [7], when packets are lost during video transmission, significant errors may appear due to the corruption of related video data. Moreover, transmission errors in one frame may also propagate to the subsequent frames since the predictive coding structures are employed [20], [21]. This problem may be even worse in the video applications where the “IPPP” coding structure is employed. In this quality evaluation module, the video coding quality and the parameters calculated by the parameter analysis module are employed to evaluate the video quality affected by the packet loss. This module is also a focus of the proposed parametric planning model.

In the following section, the proposed parameter analysis method and the video quality evaluation method for the distortion caused by packet loss will be discussed in detail.

III. EVALUATION OF DISTORTION CAUSED BY PACKET LOSS
For video streaming applications, the information of the impaired frame is very important to the video quality since the video sequence is constituted by frames. In order to accurately indicate the influence of packet loss on the video quality for IPTV services, this section first analyzes the channel characteristics, where the probability distribution of the channel states will be calculated. Considering the temporal dependence between frames, some frame- and sequence-level parameters will be presented to indicate the influence of the direct packet loss and error propagation on the video quality. A parametric planning model is finally proposed to evaluate the video quality.

A. Four-state Markov channel model analysis
With respect to the parametric planning model, the characteristics of the channel (e.g., packet loss rate, burst
packet loss rate, channel state transition probabilities) are known a priori in the planning phase. Generally, the finite-state Markov model is a good approximation of the actual packet loss processes for both wired and wireless channels [14]. However, the more elaborate division of channel states will increase the complexity of analyzing the finite-state Markov model and calculating the parameters of planning model. Therefore, it should make a trade-off between the accuracy and complexity of channel simulation [22].

According to the study in [15], the good and bad state run length distribution often tends to behave like a mixture of two geometric distributions. This characteristic just coincides with that of four-state Markov model (4SMM). Moreover, the transition probabilities in 4SMM can be established without having to run extensive physical layer simulations, thus it is comparatively easy to implement and analyze for 4SMM. Similar to the P.1201 and G.1071 models, this paper uses a 4SMM to emulate the sophisticated packet loss process [16], [23], as illustrated in Fig. 3. It is noted that in the 4SMM, two period types are involved, namely, the burst period and the gap period. The burst period, as discussed in [16], is defined as a longest sequence beginning and ending with a loss during which the number of consecutive received packets is less than a specified value (This value varies with the network and service scenarios, and an example of the suitable value for video service is 64 according to [16]). The other periods are classified as the gap periods. These two periods appear alternatively. In each period, there are two kinds of states corresponding to the fact that the packet is received and lost, respectively. Provided with the period to which the packet belongs, the four states of the 4SMM model are defined as:

1. State A: isolated packet lost in a gap period
2. State B: packet received successfully in a gap period
3. State C: packet lost in a burst period
4. State D: packet received in a burst period

It can be found from Fig. 3 that there are several rules for the transfer between different states. For example, State A can be transferred only to State B, while State B can be transferred to State A and C as well as to itself. State C can be transferred not only to State B and D, but also to itself. State D can be transferred to State C and itself, respectively.

\[
\begin{align*}
\text{State} & \quad \text{Burst Period} \quad \text{Gap Period} \quad \text{Burst Period} \\
\text{Loss Pattern} & \quad 0^{299}, 0^{11}, 000, 10^{12}, 01^{32}, 01^{155}, 111, 00, 1^{\ldots} \\
\text{State} & \quad B-B \quad C-C \quad D-D \quad C-D \quad D-D \quad C-D \quad C-D \quad C-D \quad C-D \quad \ldots
\end{align*}
\]

Fig.4. Illustration of packet loss in the four-state Markov model by a binary sequence

Given a loss pattern of the 4SMM, the burst and gap periods are detected firstly according to their definitions. Then, the state for each packet is clarified by checking the packet is received or lost in a specified period. Fig. 4 illustrates the four states by a binary sequence where the value 1 indicates that current packet is lost and the value 0 denotes that current packet is correctly received. The values of 239, 175, 132 and 155 indicate the number of consecutive 0. It is obvious that each state is easily distinguished from other states and the state transitions obey the rules specified in Fig. 3.

For a specific 4SMM, as illustrated in Fig. 3, the matrix of transition probability \( \Pi \) can be expressed as:

\[
\Pi = \begin{bmatrix}
P_{AA} & P_{AB} & P_{AC} & P_{AD} \\
P_{BA} & P_{BB} & P_{BC} & P_{BD} \\
P_{CA} & P_{CB} & P_{CC} & P_{CD} \\
P_{DA} & P_{DB} & P_{DC} & P_{DD}
\end{bmatrix} = \begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix},
\]

where \( p_{ij}, (i,j \in \{A, B, C, D\}) \) is the state transition probability between two states. It can be indicated as follows [24]:

\[
p_g = \Pr(X_g = 1|X_g = i),
\]

where \( \Pr \) is the probability, and \( X_g \) and \( X_h \) are random variables. This channel model can be characterized by the parameters \( g, h, f, i, j, k, m, n \). The values of these parameters can be determined using the maximum likelihood estimators for a sample trace [25]. Taking the parameter \( g \) as an example, it indicates the state transition probability from State B to State A. The maximum likelihood estimator of \( g \) for a sample trace is:

\[
g = \frac{n_{BA}}{n_g},
\]

where \( n_{BA} \) is the number of times in the observed time series that state B follows state A and \( n_g \) is the number of times state B occurs in the trace. The values of other channel parameters \( (h, f, i, j, k, m, n) \) can be determined using the same method.

Therefore, the channel model can be determined when the values of \( g, f, i, j, m \) are provided. For the parametric planning model, the values of these parameters are assigned in advance and can be used as an input [18], [23].

Considering that the states of 4SMM constitute an irreducible closed set, the stationary probabilities of planning model, the values of these parameters are assigned in advance and can be used as an input [18], [23].

Considering that the states of 4SMM constitute an irreducible closed set, the stationary probabilities of planning model, the values of these parameters are assigned in advance and can be used as an input [18], [23].

These stationary probabilities form the stationary probability distribution \( \pi \). It is a row vector that keeps unchanged after being applied with the operation of transition matrix \( \Pi \) [14], [24], which can be expressed as:

\[
\pi \cdot \Pi = \pi,
\]

\[
\pi = [P_A, P_B, P_C, P_D],
\]
where $P_A, P_B, P_C, P_D$ are the stationary probabilities of State A, B, C and D, respectively. They are non-negative and sum to 1. According to (2), (8), and (9), the values of $B$, $C$ and $D$, respectively. They are non-negative and sum to 1.

Given a specific video service, the average number of packets in each frame can be determined by the ratio of $B_F$ and packet size. In the rest of this subsection, these parameters will be calculated when each frame is encapsulated into one packet and into multiple packets, respectively.

1) One packet per frame

When each frame is encapsulated into one packet for transmission, the packet loss will directly lead to the frame loss. In such a case, the frame loss process is the same as the packet loss process, which can be characterized via a four-state Markov model as well.

(i) AFLF: For the video streaming suffering from packet loss, the value of AFLF $V_{AFLF}$ can be calculated using the frame loss rate and the length of GOP, which is expressed as:

$$V_{AFLF} = P_F \cdot L_G = P_F \cdot L_G = (P_A + P_C) \cdot L_G,$$  \hspace{1cm} (14)

where $L_G$ is the length of GOP, $P_F$ is the frame loss rate, $P_A$ is the packet loss rate that is equal to the sum of the random packet loss rate $P_S$ and burst packet loss rate $P_C$. Because each frame is encapsulated in one packet, $P_F = P_S$.

(ii) ENIF: When the $i^{th}$ frame in a GOP is subject to errors caused by packet loss, its subsequent frames are usually error-prone because the $i^{th}$ frame may be used as the reference. This contamination will not stop until the next synchronization point, typically an I-frame, is reached. Thus, the number of impaired frames is determined by the GOP length and the position of the lost frame.

For a video with the GOP length $L_G$, as shown in Fig. 6, there are a total of $L_G$ loss patterns to be considered. More specifically, the value 1 indicates that the frame is affected by direct packet loss or error propagation, while the value 0 means that the frame is intact and not influenced by the packet loss. Taking the loss pattern of index 2 in Fig. 6 as an example, the second frame in a GOP is suffered from packet loss, and the total number of impaired frames is $L_G - 1$. In the Markov channel, the packet loss event exhibits dependencies over time [14]. Thus, the probability of each loss pattern should be calculated using the stationary probabilities of different states and corresponding transition probabilities. Based on this analysis, the expectation of the impaired frame number $V_{ENIF}$ for a single packet loss is calculated as:

$$V_{ENIF} = \sum_{k=1}^{L_G} P_{NL}(k) \cdot N_{k}(k),$$

\hspace{1cm} (15)

where $N_{k}(k)$ and $P_{NL}(k)$ are the number of the impaired frames and the probability of the loss pattern for the $k^{th}$ index, respectively, as illustrated in Fig. 6.

When multiple packet losses occur in a GOP, the high frame loss frequency will lead to low expectation of the impaired frame number for each packet loss. As illustrated in Fig. 7, the values of $E[N_{L1}], E[N_{L2}]$ and $E[N_{L3}]$ indicate the expectations of the impaired frame numbers for different packet loss events in a GOP. Obviously, $L_G > E[N_{L1}] > E[N_{L2}] > E[N_{L3}] > ... > E[N_{Lk}]$. The relationship among these values can be expressed as follows:

$$E[N_{L1}] > E[N_{L2}] > E[N_{L3}] > ... > E[N_{Lk}].$$
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2638621, IEEE Transactions on Multimedia

Frame is encapsulated into multiple packets.

Independent and identically distributed. Next, the proposed and characterized via a Bernoulli model [26], which is [14]. In such a case, the frame loss process can be simplified and the burst packet loss concentrate on affecting individual frames, rather than spreading across multiple frames [14]. In such a case, the frame loss process can be simplified and characterized via a Bernoulli model [26], which is independent and identically distributed. Next, the proposed parameters will be calculated under the condition that each frame is encapsulated into multiple packets.

(i) AFLF: If one of the packets of a frame is lost, this frame is defined as a lossy frame. For a frame with the number of \( V_{PP} \) packets, the probability of the lossy frame can be calculated as:

\[
P_f = 1 - \left( P_g \cdot h^{\gamma_{prl}} + P_p \cdot \eta^{\gamma_{prl}} \right),
\]

where the value of \( V_{PP} \) can be calculated by the ratio of \( B_f \) and packet size. The value of \( P_g \cdot h^{\gamma_{prl}} \) indicates the probability that all packets of the current frame are in State B (error free), and the value of \( P_p \cdot \eta^{\gamma_{prl}} \) indicates the probability that all packets of the current frame are in State D (error free in a burst period). Thus, the value of \( 1 - P_g \cdot h^{\gamma_{prl}} - P_p \cdot \eta^{\gamma_{prl}} \) is the probability of a frame suffering from packet loss. Accordingly, the value of AFLF \( V_{AFLF} \) is achieved by:

\[
V_{AFLF} = P_f \cdot L_g
\]

(ii) ENIF: To calculate the value of ENIF, the probability of each loss pattern and the number of impaired frames under each loss pattern should be determined. Fig. 8 gives the probability of each loss pattern and the corresponding number of impaired frames for single packet loss. It can be found that the probability of each loss pattern is different from that in Fig. 6. This is because the probability distributions of the frame loss in these two conditions are different from each other. Based on the analysis above, the value of ENIF \( V_{ENIF} \) can be calculated by accumulating the product of the number of impaired frames for each pattern and the corresponding probability as follows:

\[
V_{ENIF} = E \left\{ L_{ni} \right\} = \sum_{i=1}^{L_g} P_f (k) \cdot N_f (k)
\]

According to the L’ Hospital’s rule [27], the value of ENIF is equal to 0 when \( P_f \) is 0. This result indicates that if there is no packet loss occurs in the video streaming, the number of the impaired frames is equal to 0.

When multiple packet losses occur in a GOP, the high frame loss frequency may also lead to low expectation of the impaired frame number for each packet loss. In such a case, the average impaired frame number of each loss should be regularized by the frame loss frequency \( V_{AFLF} \), which can refer to (18).

(iii) EIRF: When a packet of the frame is lost, as illustrated in Fig. 9, the successive packets of this frame have to be discarded because they usually cannot be decoded correctly. Therefore, the impaired ratio of the frame is closely related to
the number of the packets in a frame and the position of the packet loss occurs in the frame.

Considering the temporal dependency, the packet loss process in a frame can be characterized by a Markov model as well. Given a frame with $V_{PEF}$ packets, there are $V_{PEF}$ packet loss patterns to be considered. The expectation of the impaired ratio of the frame under each loss pattern and its probability as follows:

$$V_{EIRF} = \sum_{k=1}^{V_{PEF}} P_k(1 - h)^{V_{PEF}}(1 - h)^{PEF}$$

where $R_k(k)$ is the impaired ratio of the frame under a $k^{th}$ loss pattern. Particularly, when $V_{PEF}$ is equal to 1, the value of EIRF calculated by (22) is also equal to 1. It indicates that if each frame is encapsulated into one packet to transmit, the packet loss will lead to a whole frame loss.

### C. Video quality assessment model

Based on the analysis above, it can be found that apart from the basic statistical information (e.g., bit-rate and packet loss rate), other parameters (such as the $V_{AFLF}$, $V_{ENIF}$, and $V_{EIRF}$) are useful for reflecting the packet or frame loss behaviors and can be obtained by analyzing the characteristics of the channel and video coding. However, for the parametric planning model, how to find the relationship between these parameters and the video quality is also a key issue to be solved.

With respect to the video transmitted over the network, the video quality will be affected by the coding distortion and the distortion caused by packet loss. As mentioned in section II, the video coding quality has been estimated using the parameters such as the coded bit-rate. Here, the focus is on the evaluation of the distortion caused by packet loss. For the video streaming with packet loss, the normalized distortion caused by packet loss $D_l$ can be expressed as:

$$D_l = \frac{Q_l - Q_0}{Q - Q_1},$$

where $Q_l$ is the video quality affected by packet loss. It can be found from (23) that the value of $Q_l$ is closely related to $D_l$ when $Q_0$ is known.

In order to estimate the value of $D_l$ for a given network scenario, the relationship between $D_l$ and $V_{AFLF}$, $V_{ENIF}$, $V_{EIRF}$ should be clarified, respectively. The one-way analysis of variance (ANOVA) tests are performed to check their relationship. The corresponding F-values are 20.06, 51.91, and 15.36. All p-values of the F-test results are smaller than 0.01 at95% level, as illustrated in Table III. It indicates that the proposed parameters have a significant positive correlation with $D_l$. Here, a specific function $D_l(V_{AFLF}, V_{ENIF}, V_{EIRF})$ is defined. The monotonicity and the convergence of the function should meet the following requirements:

(A) The value of $D_l$ should increase with the raise of $V_{AFLF}$, $V_{ENIF}$, and $V_{EIRF}$, respectively. This is because the larger values of $V_{AFLF}$, $V_{ENIF}$, and $V_{EIRF}$ indicate more packet loss, more frames affected by error propagation, and severer error in each frame, respectively.

(B) If there is no packet loss in the transmission, the values of $V_{AFLF}$, $V_{ENIF}$, and $V_{EIRF}$ are all equal to 0. In such a case, the value of $D_l$ is also equal to 0.

(C) If the values of $V_{AFLF}$, $V_{ENIF}$, and $V_{EIRF}$ are considerably large, the value of $D_l$ should be approximated to the maximum value 1.

Correspondingly, an evaluation model for the video distortion caused by packet loss is proposed to satisfy the above constraints:

$$D_l = 1 - \exp(-v_1 \cdot V_{AFLF}^{v_6} \cdot V_{ENIF}^{v_7} \cdot V_{EIRF}^{v_8}),$$

where $v_1$, $v_6$, $v_7$, and $v_8$ are empirical parameters obtained by regression. Submitting (23) into (22), the video quality affected by packet loss can be calculated as follow:

$$Q_l = 1 + (Q_0 - 1) \cdot \exp(-v_1 \cdot V_{AFLF}^{v_6} \cdot V_{ENIF}^{v_7} \cdot V_{EIRF}^{v_8}).$$

Until now, the parametric planning model has been established, where the channel and video characteristics are combined together to evaluate the video quality. Table IV lists
all the notations of this section. Considering the value of $Q_c$ is estimated by $B_R$ and $F_R$ in (1), there are five parameters in the proposed parametric planning model, namely $B_S$, $F_S$, $V_{\text{AVL}}$, $V_{\text{ENS}}$, and $V_{\text{ERR}}$. The parameters $V_{\text{AVL}}, V_{\text{ENS}}$, and $V_{\text{ERR}}$ can be further estimated from the probability distribution of the channel states (state transition probability). It considers the burst loss, the error propagation and the correlation between individual packet loss events.

### IV. EXPERIMENTAL SETTINGS AND RESULTS

The experiments in this paper consist of two sessions: the training session and the validation session. For the training session, it mainly focuses on training the coefficients of the proposed model. For the validation session, it is used to check the performance of the proposed planning model.

#### A. Experiment settings for training

In the training session, four standard video sequences with different contents were employed. The involved sequences included: CrowdRun, DucksTakeOff, Stockholm, and BasketballDrive. Each video sequence with length of 10 seconds has three resolutions: QVGA, HVGA and 720P. All sequences were compressed by the FFmpeg 0.4.9 codec with x.264 library [28]. The coding structure was “IPPP”. The number of slices per frame was 1. The length of GOP was 60 and the number of reference frames was 1. The detailed coding information such as the coded bit-rate and the frame rate are listed in Table V.

To simulate the packet loss process in IP networks, the compressed bit streams were packetized following the format defined by the RTP payload format for H.264 [29], where the maximum size of the packet was set as 1500 bytes. A four-state Markov model was employed to simulate the packet loss distribution. The channel state transition probabilities $i, j, k, m, n$ were set at 0.3, 0.65, 0.05, 0.25, 0.75, respectively, which were recommended by ITU Study Group 12 for network video services [23]. The value of $f$ and $g$ led to different packet loss rates. Here, the values of $f$ and $g$ were set at 0.0012, 0.0023, 0.0047, 0.0072 and 0.0122, which correspond to the packet loss rates 0.5%, 1%, 2%, 3% and 5%, and the burst packet loss 0.15%, 0.5%, 0.6%, 0.9% and 1.5%, respectively, as illustrated in Table V. All sequences were decoded using H.264 decoder FFmpeg 0.4.9 with the zero motion error concealment techniques. When a packet in a frame was lost, the successive packets of this frame were all discarded as well. To enable decoding, the first packet (in the instantaneous data refresh frame) and the last packet (the loss of this packet cannot be detected by the decoder) were not dropped in the experiment.

#### TABLE IV

**NOTATIONS FOR THE PARAMETRIC PLANNING MODEL.**

<table>
<thead>
<tr>
<th>Notations</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Pi$</td>
<td>transition matrix</td>
</tr>
<tr>
<td>$\pi$</td>
<td>stationary probability distribution</td>
</tr>
<tr>
<td>$L_G$</td>
<td>length of GOP</td>
</tr>
<tr>
<td>$V_{\text{avf}}$</td>
<td>average number of packets in a frame</td>
</tr>
<tr>
<td>$V_{\text{avr}}$</td>
<td>average frame loss frequency</td>
</tr>
<tr>
<td>$V_{\text{ENS}}$</td>
<td>expectation of the number of impaired frames</td>
</tr>
<tr>
<td>$V_{\text{ERR}}$</td>
<td>expectation of the impaired ratio of a frame</td>
</tr>
<tr>
<td>$D_{ij}$</td>
<td>normalized distortion caused by packet loss</td>
</tr>
<tr>
<td>$\hat{Q}$</td>
<td>video quality affected by packet loss</td>
</tr>
<tr>
<td>$g, h, f, i, j, k, m, n$</td>
<td>state transition probability</td>
</tr>
</tbody>
</table>

#### TABLE V

**EXPERIMENTAL SETTINGS FOR THE TRAINING SET**

<table>
<thead>
<tr>
<th>Index</th>
<th>Bitrate (kbps)</th>
<th>Frame rate</th>
<th>Packet loss rate</th>
<th>Burst packet loss rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>QVGA</td>
<td>128, 192</td>
<td>15</td>
<td>0, 0.5%</td>
<td>0, 0.15%</td>
</tr>
<tr>
<td></td>
<td>384, 768</td>
<td>30</td>
<td>1%, 2%, 3%</td>
<td>0.3%, 0.6%</td>
</tr>
<tr>
<td>HVGA</td>
<td>512, 768, 1536</td>
<td>30</td>
<td>3%, 5%</td>
<td>0.9%, 1.5%</td>
</tr>
<tr>
<td>720P</td>
<td>512, 1024, 2048, 4096</td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In order to obtain the video quality, a subjective test was carried out following the guide-lines specified by ITU-T recommendation P.913 [30]. The test environment was controlled, where the room luminosity was between 100 and 200 Lux and the noise of the laboratory was 40-50 dB. The monitor used for display was 22-inch LCD Flat Panel, with a resolution of 1920 $\times$ 1080 pixels. The videos were displayed at their native resolution to prevent any distortions caused by scaling operations. The viewing distance was set between 4H-5H (H is the picture height).

A total of 25 non-expert viewers participated in this subjective test, including 12 females and 13 males. All participants were university students aged between 23 and 28 years, and they were screened for visual acuity and color blindness. Subjective video quality was assessed using a single-stimulus presentation method and a 5-point absolute category rating (ACR) scale [30]. The duration of the subjective test was limited to 30 minutes to prevent eye strain and fatigue. Before the formal test, the views were asked to...
watch four examples to get familiar with the rating process. During the formal test, the users were instructed to watch each video clip once, and to rate the video quality immediately after watching it. All videos in the experiment were viewed by each subject. After the test, the users were screened with regard to the accuracy of their rating values. The standard exclusion procedures were followed as specified in [31]. After this screening process, the rating samples of 2 out of 25 subjects (8%) were discarded. The video quality of each video sequence was finally obtained by averaging all the 23 subject’s rating scores, also known as the mean opinion score (MOS) [30]. There were 288 video quality scores for 288 video clips and all these test data constituted the training data set TR.

Operational coefficients given in Table VI were trained using the data in TR. The values of $v_1$, $v_2$, $v_3$ and $v_4$ were obtained by training the coded data without packet loss in TR, in accordance with the method proposed in [19]. More specifically, the values of $v_1$, $v_2$, $v_3$ were fitted using the coded data in TR with the frame rate of 30fps by the least square error fitting. Then, the value of $v_4$ was fitted using the coded data with frame rate of 15 fps. Considering the effect of packet loss, the value of $D_1$, for the data with packet loss in TR was calculated according to (23). Then, the coefficients $v_1$ to $v_4$ were trained using the values of $D_1$, $V_{RES}$, $V_{END}$, and $V_{ERF}$ by the least square error fitting. It should be noted that the coefficients under different video resolutions should be trained separately.

### TABLE VI

**COEFFICIENTS IN THE PROPOSED METHOD**

<table>
<thead>
<tr>
<th>Index</th>
<th>$v_1$</th>
<th>$v_2$</th>
<th>$v_3$</th>
<th>$v_4$</th>
<th>$R^2$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>QVGA</td>
<td>3.75</td>
<td>1.07</td>
<td>2.36</td>
<td>0.20</td>
<td>0.723</td>
<td>0.363</td>
</tr>
<tr>
<td>HVGA</td>
<td>3.79</td>
<td>1.11</td>
<td>2.17</td>
<td>0.21</td>
<td>0.748</td>
<td>0.352</td>
</tr>
<tr>
<td>720P</td>
<td>3.82</td>
<td>1.16</td>
<td>2.04</td>
<td>0.25</td>
<td>0.763</td>
<td>0.348</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Index</th>
<th>$v_1$</th>
<th>$v_2$</th>
<th>$v_3$</th>
<th>$v_4$</th>
<th>$R^2$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>QVGA</td>
<td>0.32</td>
<td>1.21</td>
<td>0.04</td>
<td>1.69</td>
<td>0.630</td>
<td>0.561</td>
</tr>
<tr>
<td>HVGA</td>
<td>0.26</td>
<td>0.96</td>
<td>0.04</td>
<td>1.52</td>
<td>0.664</td>
<td>0.542</td>
</tr>
<tr>
<td>720P</td>
<td>0.72</td>
<td>1.23</td>
<td>0.03</td>
<td>2.21</td>
<td>0.688</td>
<td>0.538</td>
</tr>
</tbody>
</table>

### B. Experimental settings for validation

In the validation session, the performance of the proposed method was checked using six sequences including: ParkRun, Riverbed, IntoTree, Sunflower, ParkJoy, and Touchdown. The detail experiment settings are listed in Table VII. The procedures of the subjective test were the same as those in the training test. All the subjective rating values constituted the validation data set VL.

### TABLE VII

**EXPERIMENTAL SETTINGS FOR VALIDATION SET**

<table>
<thead>
<tr>
<th>Index</th>
<th>Bitrate (kbps)</th>
<th>Frame rate</th>
<th>Packet loss rate</th>
<th>Burst packet loss rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>QVGA</td>
<td>96, 128</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HVGA</td>
<td>256, 512</td>
<td>30</td>
<td>0, 0.5%, 0, 0.15%</td>
<td>0.1%, 2%, 0.3%, 0.6%</td>
</tr>
<tr>
<td>720P</td>
<td>384, 512, 832, 1600</td>
<td>30</td>
<td>3%, 5%, 0.9%, 1.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>512, 768, 1536, 3840</td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Three metrics suggested by VQEG [32] were employed in this work. They were Pearson correlation coefficient (PCC) for linearity, root-mean-squared error (RMSE) for accuracy, and Spearman rank order correlation coefficient (SROCC) for monotonicity. The maximum value of PCC and SROCC is 1. Generally, the smaller of the RMSE and the larger of the PCC and SROCC indicate the better performance of the model.

The proposed model was compared with other three parametric planning models [9], [10], [12]. All these models consider the effect of burst packet loss for video quality evaluation. Because the video resolutions are not included in these models, the corresponding model coefficients were re-trained on the proposed TR dataset, according to their recommended methods in [9], [10] and [12], respectively. Moreover, for simplicity, the models in [9], [10], and [12] are denoted as VQA1, VQA2 and VQA3, respectively.

Table VIII compares the performance of each model. It is observed that the proposed model has larger PCC and SROCC values and smaller RMSE values than other models under various resolutions. Fig. 10 visualizes the performance of each model by using the scatter plots of the objective and subjective scores in VL dataset. There is a better linear relationship between the predicted video quality by the proposed model and MOS, which indicates that the video quality predicted by the proposed model is more close to the MOS.

To further check the prediction accuracy of the model, the statistical significance analysis and hypothesis test were performed. Firstly, the assumption of Gaussianity of the scores estimated by each objective VQA model was checked using the Kolmogorov-Smirnov test (K-S test). According to the test results by SPSS 17.0 [33] (scores for each resolution, p>0.05), the null hypothesis (the scores have a standard normal distribution) could not be reject at the 5% level for any models. Therefore, the assumption of Gaussianity was valid for the scores estimated by all models.

Then, an F-test based on the residuals between MOS and the video quality estimated by different objective VQA models was performed to statistically compare performance of these models. The null hypothesis was that the variance of residuals from one
However, the average standard deviations of PCC, RMSE and SROCC are 0.004, 0.007 and 0.003, respectively. The standard deviations of PCC, SROCC and RMSE are comparatively small when the channel model parameters slightly changes. Thus, the performance of the proposed model has certain robustness to the accuracy of the channel model parameters.

$$\text{Null Hypothesis } \iff \sigma^2_{\text{MOS} - \text{MOS}^1} = \sigma^2_{\text{MOS} - \text{MOS}^2}$$  
(26)

where MOSp1 and MOSp2 were the video quality scores estimated by two different VQA models. $\sigma^2_{\text{MOS} - \text{MOS}^p}$ indicates that the variance of residuals between MOS and MOSp. The results of the statistical significance test are presented in Table IX, where the values are the probabilities when the null hypothesis of equal variances is not rejected. According to the results in Tables VIII and IX, the proposed model demonstrates a significantly superior prediction performance to other models. It can efficiently estimate the video quality for IPTV services which suffered from packet loss.

### TABLE VIII

**PERFORMANCE OF EACH MODEL**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PCC</td>
<td>0.809</td>
<td>0.714</td>
<td>0.733</td>
<td>0.760</td>
</tr>
<tr>
<td>QVGA</td>
<td>RMSE</td>
<td>0.521</td>
<td>0.592</td>
<td>0.584</td>
</tr>
<tr>
<td></td>
<td>SROCC</td>
<td>0.755</td>
<td>0.643</td>
<td>0.665</td>
</tr>
<tr>
<td>HVGA</td>
<td>PCC</td>
<td>0.824</td>
<td>0.814</td>
<td>0.787</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.511</td>
<td>0.556</td>
<td>0.552</td>
</tr>
<tr>
<td></td>
<td>SROCC</td>
<td>0.816</td>
<td>0.805</td>
<td>0.771</td>
</tr>
<tr>
<td>720P</td>
<td>PCC</td>
<td>0.850</td>
<td>0.821</td>
<td>0.834</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.546</td>
<td>0.615</td>
<td>0.605</td>
</tr>
<tr>
<td></td>
<td>SROCC</td>
<td>0.861</td>
<td>0.785</td>
<td>0.786</td>
</tr>
<tr>
<td>ALL</td>
<td>PCC</td>
<td>0.828</td>
<td>0.787</td>
<td>0.786</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.519</td>
<td>0.588</td>
<td>0.581</td>
</tr>
<tr>
<td></td>
<td>SROCC</td>
<td>0.821</td>
<td>0.789</td>
<td>0.771</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed Model</td>
<td>1.000</td>
<td>0.027</td>
<td>0.033</td>
</tr>
<tr>
<td>Yamagishi [9]</td>
<td>0.027</td>
<td>1.000</td>
<td>0.119</td>
</tr>
<tr>
<td>Garcia [10]</td>
<td>0.033</td>
<td>0.119</td>
<td>1.000</td>
</tr>
<tr>
<td>G.1071 [12]</td>
<td>0.041</td>
<td>0.087</td>
<td>0.052</td>
</tr>
</tbody>
</table>

Considering that there may be a deviation between the actual network parameters and their statistical values, the sensitivity of the proposed parametric planning model to the accuracy of the channel parameters was also checked. Here, the deviations of the channel parameters i, j, m were set at -10%, -5%, 0%, 5% and 10%, respectively, as illustrated in Table X. It can be found that, when the deviation of channel model parameters is in a range from -10% to +10%, the maximum errors (ME) of PCC, RMSE and SROCC between MOS and the predicted video quality are 0.7%, 2.7% and 0.7% (within 95% level), respectively. The maximum standard deviations (MSD) of PCC, RMSE and SROCC are 0.004, 0.007 and 0.003, respectively. However, the average standard deviations of PCC, RMSE and SROCC between the predicted video quality by different methods and MOS are 0.02, 0.032 and 0.021, respectively. The standard deviations of PCC, SROCC and RMSE are comparatively small when the channel model parameters slightly changes. Thus, the performance of the proposed model has certain robustness to the accuracy of the channel model parameters.

### TABLE IX

**STATISTICAL SIGNIFICANCE MATRIX BASED ON RESIDUAL BETWEEN MOS AND THE PREDICTED VIDEO QUALITY.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PCC</td>
<td>0.805</td>
<td>0.807</td>
<td>0.812</td>
<td>0.814</td>
</tr>
<tr>
<td>QVGA</td>
<td>RMSE</td>
<td>0.513</td>
<td>0.517</td>
<td>0.523</td>
</tr>
<tr>
<td></td>
<td>SROCC</td>
<td>0.751</td>
<td>0.752</td>
<td>0.756</td>
</tr>
<tr>
<td>HVGA</td>
<td>PCC</td>
<td>0.828</td>
<td>0.827</td>
<td>0.823</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.524</td>
<td>0.516</td>
<td>0.506</td>
</tr>
<tr>
<td></td>
<td>SROCC</td>
<td>0.822</td>
<td>0.820</td>
<td>0.815</td>
</tr>
<tr>
<td>720P</td>
<td>PCC</td>
<td>0.846</td>
<td>0.849</td>
<td>0.852</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.561</td>
<td>0.551</td>
<td>0.544</td>
</tr>
<tr>
<td></td>
<td>SROCC</td>
<td>0.864</td>
<td>0.862</td>
<td>0.858</td>
</tr>
</tbody>
</table>

### V. CONCLUSIONS

Because the actual video streams cannot be obtained during the planning phase in practice, only little information can be used in the parametric planning model. Consequently, how to accurately estimate the video quality is rather challenging. In this paper, a parametric planning model was proposed for IPTV services. The principal contributions of this work are to check the necessity of combining the channel and video characteristics to evaluate the video quality in the planning phase. Considering the fact that the burst packet loss and the temporal dependence between frames will lead to nonrandom frame distortion, several sequence-level and frame-level parameters for video quality evaluation have been derived from the channel models. Using these parameters, the proposed planning model can better indicate the effects of direct packet loss and error propagation. Experimental results have demonstrated that the proposed model has excellent performance in video quality evaluation and it can be used as an effective tool for network or QoE planning.

It should be noted that the video quality may also be affected by other factors such as different encoding structure, frame or slice type, error control and concealment techniques. Our future work will consider the influence of these factors on the video quality assessment. Moreover, the proposed model can be further extended to interactive video services by taking other factors into consideration, such as the one-way delay and jitter.

REFERENCES


Demonstrated that

\[ \Pi = \begin{bmatrix}
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix} \]

\[ \pi = [P_A, P_B, P_C, P_D] \]

The equation can be expressed as follows

\[ P_{11} P_{22} + P_{12} P_{21} + P_{13} P_{31} + P_{14} P_{41} = P_A \]

According to (2), there are

\[ P_{11} P_{22} P_{33} P_{44} P_{55} + P_{12} P_{21} P_{32} P_{41} P_{54} + P_{13} P_{31} P_{32} P_{43} P_{54} + P_{14} P_{41} P_{42} P_{43} P_{54} = P_D \]

Take g, i, j, k, m and n into (4), we can obtain that

\[ P_{11} P_{22} + P_{12} P_{21} + P_{13} P_{31} + P_{14} P_{41} + P_{22} P_{33} P_{44} P_{55} + P_{21} P_{32} P_{41} P_{54} + P_{23} P_{31} P_{32} P_{43} P_{54} + P_{24} P_{41} P_{42} P_{43} P_{54} = P_D \]

Considering that

\[ h = 1 - f - g \]

\[ k = 1 - i - j \]

\[ n = 1 - m \]

\[ P_A + P_B + P_C + P_D = 1 \]
Take (6) into (5) and solve equations, we can obtain that

\[ P_s = \frac{mgj}{(m+k) \cdot f + (1+g) \cdot mj} \]
\[ P_b = \frac{mj}{(m+k) \cdot f + (1+g) \cdot mj} \]
\[ P_m = \frac{mjf}{(m+k) \cdot f + (1+g) \cdot mj} \]
\[ P_d = \frac{f}{(m+k) \cdot f + (1+g) \cdot mj} \]

To calculate the values of \( V_{ENF} \) and \( V_{EBFS} \), there are two cases to be considered.

**Case 1:** When each frame is encapsulated into one packet for transmission, the packet loss will directly lead to the frame loss. In such a case, the value of \( V_{ENF} \) can be derived as follows:

**Proposition 2:**

\[ V_{ENF} = E \{ N_k \} = \sum_{k=1}^{L_{G}} P_k(k) \cdot N_k(k) \]
\[ = (L_{G} - P_a(1 - h^{n-1}) - P_d(1 - n^{r-1})) \cdot (1 - P_bh^{n-1} - P_db^{n-1})^{-1} \]

**Proof:**

\[ V_{ENF} = E \{ N_k \} = \sum_{k=1}^{L_{G}} P_k(k) \cdot N_k(k) \]
\[ = P_a(1) \cdot N_1(1) + P_a(2) \cdot N_1(2) + P_a(3) \cdot N_1(3) + \ldots + P_a(L_{G}) \cdot N_1(L_{G}) \]
\[ - \left( P_a + P_b(1 - h^{n-1}) - P_d(1 - n^{r-1}) \right) \cdot L_{G} \]
\[ + P_b(1 + P_d) \cdot h^{n-1} \cdot P_d \cdot n^{r-1} \cdot m \cdot (L_{G} - 2) \]
\[ + \ldots + \left( P_a + P_b(1 - h^{n-1}) - P_d(1 - n^{r-1}) \right) \cdot L_{G} \]
\[ - \left( P_a + P_b(1 - h^{n-1}) - P_d(1 - n^{r-1}) \right) \cdot L_{G} \]
\[ = \frac{(P_a + P_b)L_{G} + P_a(g + f)
\sum_{i=1}^{L_{G}} h^i(L_{G} - i) + P_m\sum_{i=1}^{L_{G}} n^r(L_{G} - r)}{1 - P_bh^{n-1} - P_db^{n-1}} \]

Let

\[ S_{a1} = \sum_{i=1}^{L_{G} - 1} h^i(L_{G} - i), \quad S_{a2} = \sum_{i=1}^{L_{G} - 1} n^r(L_{G} - r) \]

There is

\[ S_{a1} = \sum_{i=1}^{L_{G} - 1} h^i(L_{G} - i) \]
\[ = (L_{G} - 1)h(L_{G} - 2) + h^2(L_{G} - 3) + \ldots + h^{L_{G}-2} \]
\[ = h(L_{G} - 1) + h^2(L_{G} - 2) + h^3(L_{G} - 3) + \ldots + h^{L_{G}-1} \]

Subtract (2) from (3), we can obtain that

\[ (h-1) \cdot S_{a1} = -(L_{G} - 1) + h + h^2 + \ldots + h^{L_{G}-2} + h^{L_{G}-1} = -L_{G} + \frac{1 - h^{L_{G}}}{1 - h} \]
\[ S_{a1} = \frac{L_{G} - 1 - h^{L_{G}}}{1 - h} \]

Considering that \( S_{a1} \) and \( S_{a2} \) have the same form, therefore \( S_{a2} \) can be derived following the above steps as well

\[ S_{a2} = \frac{L_{G} - 1 - n^{L_{G}}}{1 - n} \]

Take (4) and (5) into (1), it can obtain that

\[ V_{ENF} = \frac{(P_a + P_b)L_{G} + P_a(g + f)}{1 - P_bh^{n-1} - P_db^{n-1}} \]
\[ = \left( P_a + P_b \right) L_{G} + \frac{P_a(g + f)}{1 - h^{n-1}} \]
\[ + \frac{L_{G} - 1 - n^{r-1}}{1 - h^{n-1}} \]

Because \( h = 1 - f - g \), and \( n = 1 - m \)

\[ V_{ENF} = \left( P_a + P_b \right) L_{G} + \frac{P_a(g + f)}{1 - h^{n-1}} \]
\[ + \frac{L_{G} - 1 - n^{r-1}}{1 - h^{n-1}} \]

\[ = \left( P_a + P_b \right) L_{G} + \frac{P_a(g + f)}{1 - h^{n-1}} \]
\[ + \frac{L_{G} - 1 - n^{r-1}}{1 - h^{n-1}} \]

**Case 2:** When each frame is encapsulated into multiple packets for transmission, the values of \( V_{ENF} \) and \( V_{EBFS} \) can be derived as follows:

**Proposition 3:**

\[ V_{ENF} = E \{ N_k \} = \sum_{k=1}^{L_{G}} P_k(k) \cdot N_k(k) \]
\[ = \sum_{k=1}^{L_{G} - i + 1} \left( P_a(1 - P_b)^{y_k} + \frac{L_{G} - 1 - P_b^{y_k}}{1 - (1 - P_b)^{y_k}} - \frac{P_a}{P_b} \right) \]

**Proof:**

\[ V_{ENF} = E \{ N_k \} = \sum_{k=1}^{L_{G}} P_k(k) \cdot N_k(k) \]
\[ = P_a(1) \cdot N_1(1) + P_a(2) \cdot N_1(2) + P_a(3) \cdot N_1(3) + \ldots + P_a(L_{G}) \cdot N_1(L_{G}) \]
\[ - \left( P_a + P_b(1 - h^{n-1}) - P_d(1 - n^{r-1}) \right) \cdot L_{G} \]
\[ + P_b(1 + P_d) \cdot h^{n-1} \cdot P_d \cdot n^{r-1} \cdot m \cdot (L_{G} - 2) \]
\[ + \ldots + \left( P_a + P_b(1 - h^{n-1}) - P_d(1 - n^{r-1}) \right) \cdot L_{G} \]
\[ - \left( P_a + P_b(1 - h^{n-1}) - P_d(1 - n^{r-1}) \right) \cdot L_{G} \]
\[ = \frac{(P_a + P_b)L_{G} + P_a(g + f)
\sum_{i=1}^{L_{G} - 1} h^i(L_{G} - i) + P_m\sum_{i=1}^{L_{G} - 1} n^r(L_{G} - r)}{1 - P_bh^{n-1} - P_db^{n-1}} \]

Let
Subtract (2) from (3), it can obtain that

\[ S_a = \sum_{i=0}^{L_x} (L_x - i + 1) (1 - P_r)^i \]

Therefore

\[ S_a = L_x + (L_x - 1) (1 - P_r) + (L_x - 2) (1 - P_r)^2 + \cdots + (1 - P_r)^{L_x-1} \]

\[ (1 - P_r) S_a = L_x (1 - P_r) + (L_x - 1) (1 - P_r)^2 + (L_x - 2) (1 - P_r)^3 + \cdots + (1 - P_r)^{L_x} \]

\[ +2 (1 - P_r)^{L_x-1} + (1 - P_r)^{L_x} \]

\[ \Rightarrow S_a = \frac{L_x - 1}{1 - P_r} - \frac{2}{P_r} + \frac{1 - P_r}{P_r} \]

\[ \therefore \]

Subtract (2) from (3), it can obtain that

\[ S_a = \frac{L_x - 1}{1 - P_r} - \frac{2}{P_r} + \frac{1 - P_r}{P_r} \]

There is

\[ S_a = \frac{L_x - 1}{1 - P_r} - \frac{2}{P_r} + \frac{1 - P_r}{P_r} \]

Considering that \( S_{a1} \) and \( S_{a2} \) have the same form, therefore \( S_{a2} \) can be derived following the above steps as well

\[ S_{a2} = \frac{V_{prf} 1 - n^{\gamma_{prf}}}{1 - n} \]

Take (4) and (5) into (1), it can obtain that

\[ V_{EIRF} = \frac{P_e + \frac{P_s (g + f) S_{a1} + P_s m S_{a2}}{V_{prf} - 1 - P_g h^{\gamma_{prf}} - P_m h^{\gamma_{prf}}}}{1 - P_g h^{\gamma_{prf}} - P_m h^{\gamma_{prf}} - 1} \]

Because \( h = 1 - f \), and \( n = 1 - m \)

\[ V_{EIRF} = \frac{P_e + \frac{P_s (g + f) S_{a1} + P_s m S_{a2}}{V_{prf} - 1 - P_g h^{\gamma_{prf}} - P_m h^{\gamma_{prf}}}}{1 - P_g h^{\gamma_{prf}} - P_m h^{\gamma_{prf}} - 1} \]

\[ \Rightarrow S_{a1} = \sum_{i=1}^{V_{prf}} h^{i-1} (V_{prf} - i) \]

\[ S_{a2} = \sum_{i=1}^{V_{prf}} n^{i-1} (V_{prf} - r) \]

There is

\[ S_{a1} = \sum_{i=1}^{V_{prf}} h^{i-1} (V_{prf} - i) \]

\[ (h-1) S_{a1} = -(V_{prf} - 1) + h^2 (V_{prf} - 3) + \cdots + h^{\gamma_{prf}} \]

\[ s_{a2} = \frac{V_{prf} 1 - n^{\gamma_{prf}}}{1 - n} \]

\[ \Rightarrow S_{a2} = \sum_{i=1}^{V_{prf}} h^{i-1} (V_{prf} - i) \]

\[ h \times S_{a1} = \sum_{i=1}^{V_{prf}} h (V_{prf} - i) \]

\[ h (V_{prf} - 1) + h^2 (V_{prf} - 2) + \cdots + h^{\gamma_{prf}} \]

\[ \Rightarrow \]

Subtract (2) from (3), it can obtain that

\[ \Rightarrow \]