Binomial Probability Redundancy Strategy for Multimedia Transmission

Lizhuo ZHANG\textsuperscript{1,2}, Weijia JIA\textsuperscript{1,2}, Shifei ZHOU\textsuperscript{1,2}

1. School of Information Science & Engineering, Central South University, Changsha, China
2. Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong
zhanglizhuo@gmail.com; itjia@cityu.edu.hk; haiouzsf@gmail.com

Abstract—This paper proposes a Binomial Probability Redundancy Model (BPRM). This model is based on the Forward Error Correction (FEC) Reed-Solomon Coding Technique and the Interweaving Packet Loss Recovery Technique. It calculates the number of redundant packets according to the feedback from the receiver and dynamically generates these superfluous packets. Different from the Linear Probability Redundancy Model (LPRM) proposed by McKinley et al., we adopt the interweaving technique to transmit these redundant packets. We also present a bandwidth control strategy to improve quality of video data. The experimental result shows that BPRM is able to generate enough redundant packets even under congestive network condition. These packets ensure the receiver to restore the original data and achieve a better video quality compared with LPRM.

Keywords; multimedia communication; VVoIP; redundant coding; binomial probability

I. INTRODUCTION

A. Motivation

The quality of VVoIP applications will be downgraded when the traffic of Internet is crowded. Packet Loss Concealment (PLC) \cite{1} is currently a broadly used approach to solve this problem. PLC consists of sender-based packet loss recovery and receiver-based error concealment \cite{2}. In sender-based packet loss recovery, the lost packets can be recovered by the Forward Error Correction (FEC) \cite{3} redundant information generated from sender and sent to receiver. In receiver-based error concealment, the missing packets are concealed through packet insertion, interpolation or regeneration.

A lot of research has been done based on packet loss recovery strategy, especially FEC. Villalon, et al. \cite{4} has proposed a cross-layer method to adaptively multi-cast video data. And, \cite{5,6} presented an adaptive FEC mechanism, which is an agent-based reliable multicast transmission strategy. Besides, \cite{7} has implemented experiments on the Internet. Their strategy was able to adaptively adjust the number of redundant packets to avoid worsening network condition. Nevertheless, all these works are based on video transmission and do not pay attention to the voice transmission. Therefore, our work which are based on both video and voice transmission will show a great significance and innovation.

Transmission of large size packets will not waste much of bandwidth overhead. Conversely, if we use the same method to transmit voice data and their redundant packets, a significant number of bandwidth resources will be wasted. As is known, the overhead of an RTP packet is 40 bytes. The length of an audio data packet which is encoded by AMR under MR_475 mode is 13 bytes. Then, the efficiency of transmitting an audio packet can be calculated as following:

\[
\frac{13}{13+40} = 24.5\%
\]

Obviously, (1-24.5%) = 75.5% of the protocol overhead is wasted. Moreover, if the redundant packets are lost during propagation, it will greatly deteriorate the quality of playing.

B. Contributions

In this paper, we present a binomial probability redundancy model for both video and audio transmission. Based on the packet loss information from receiver side, the sender side dynamically generates the number of redundant packets from the raw data. The video redundant packets will be marked with a sequence number and sent out. Our main contributions are listed as followings:

(1) High Loss Tolerance. We partition audio redundant packets into several small pieces and pack these pieces to the end of original packets. When these packets are received at the other side, the original data will be recovered from the redundant data of the other packets if some have been lost. As a result, our model will have high tolerance ability to the packet loss.

(2) Smart Partition Technique. We use a smart partition technique, called interweaving technique, to partition the packets to ensure the pieces distributed into each packet uniformly. \cite{9} \cite{10} \cite{11} proposed a media-specific FEC sender-based repair method. This method transmits one unit of audio in the next packet. When one of these packets is lost, another packet containing the same unit will recover the loss. However, it can only cover short periods of loss because of the crude nature of the measurement. Accordingly, we introduce a useful technique called Interweaving \cite{12} for reducing the effects of loss. The redundant packets are divided into several small units...
and re-sequenced. Finally, these units are packed into the original packets and transmitted to the other peer.

(3) Efficient Bandwidth Control. In our prototype system, we dynamically adjust the quality of video data to efficiently control bandwidth consumption. When the Internet is congestive, we lower the quality to control the sending rate and avoid network congestion. Conversely, if the packet loss rate is small and Internet is not crowded, we improve the quality of video data so that the receiver can play the video data smoothly with good pictures.

II. RELATED WORK

Although there are many receiver-based strategies, we mainly focus on sender-based FEC strategies in this paper. Hence, the raw data should be processed before sent out to ensure that the raw data could be recovered correctly by using FEC technique. Meanwhile, the sender will dynamically adjust the number of redundant packets according to the feedback information from receiver. We will describe how to adjust the number of redundant packets in the followings.

A. Forward Error Correction

FEC [13] is a systemic error control for data transmission. It is not tightly dependent on the original data but can recover the lost packet correctly. However, when the lost packets are consecutive, the performance of FEC will become poor, especially in the case of continuous loss of large-size packets. Hence, the technique of interweaving is proposed. Interweaving is not a real PLC technology as it can not recover the lost packet. But it can cut down the effect of packet loss. Interweaving partitions the original data packets into many small pieces and reconstructs packets from these small units so that each new packet contains parts of the original data. As a result, the play-out of these original data can be improved greatly. Moreover, it does not create extra information. However, if a large number of packets are lost, there will be a jitter phenomenon. For the purpose of solving this problem, we will describe the way to use interweaving technique in the transmission of video and voice in section IV.

B. Reed-Solomon Code

We use Reed-Solomon Code Algorithm to handle coding error. As there are many uncertain factors over the wireless networks, this will result in error code in data packets when the packets are transmitted. Once the receiver detects there are error codes in the incoming packets, the packets will be discarded by receiver. In this case, the packet loss will become more common. To lower this negative effect of wireless, we employ FEC Reed-Solomon (RS) Coder. RS code is one of most important array of linear block codes in the field of error control. It owns strong ability of correcting instant and random error code. In our system, the sender generates redundant packets using RS and sends to the other peer. We dynamically adjust the input parameter pair \((n, k)\), where \(n\) stands for the packet number of original data and \(k\) stands for the redundant number by using FEC RS, to produce enough packets so that the receiver can decode the raw data.

C. Redundancy Transmission

To improve the quality of multimedia transmission, McKinley and Ge[15] proposed an agent-based adaptive FEC reliable transmission strategy. This strategy used Linear Probability Redundancy Model to send back the packet loss rate. The sender dynamically adjusted the number of parity packets according this feedback. However, this work takes only video transmission into consideration. To make an improvement, we proposed a Binomial Probability Model to transmit both video and audio packets.

III. BINOMIAL PROBABILITY REDUNDANCY MODEL

Although FEC is able to produce superabundance information and decode this information, Chua et al. [14] presented that there is still a defect in FEC. That is, when \(n\) packets are lost on the Internet, FEC require the sender side to produce \(n\) copies of these packets. If the lost packet number is larger than \(n\), then the rest packets will not be played out. Accordingly, they proposed ReD streaming strategy. Except the original data stream, they simultaneously resent a few copy of these streams. The receiver will selectively choose one of these resending streams to recover the lost data. Obviously, these strategies can help FEC work efficiently. But a number of network bandwidth resources are wasted. Based on the above related works, we proposed Binomial Probability Redundancy Model (BPRM). BPRM using the FEC RS coder and adjust the unit doublet value of \((n, k)\) to produce redundant packets.

A. Probability-based Transmission Model

Normally, some packets will be lost when transmitting on the Internet, especially the video RTP packet with large size. When these packets are routing in the networks, the routers will discard the large packets according to the networks congestion condition. Therefore, in order to make sure most of the large packets sent to destination, we propose to divide the large packets into several small units with the standard size of MTU. Generally, I frame of video data is very large. The average size of I frame after encoded by H.264 is 5 to 10k bytes. We then divide it to several units and generate redundant packets from these units. Finally, we send all these packets to the receiver.

1) Linear Probability Redundancy Model

In order to generate a proper number of packets, [15] proposed a Linear Probability Redundancy Model (LPRM). The number of packets sent to receiver side is \(k(1+\alpha)\). They claimed that if \(\alpha k\) packets are received, the receiver will then be able to decode the original data using FEC. Therefore, in our system, we assume that I frame are divided into \(n\) packets with size of MTU. If the last packet is smaller than MTU, the rest bytes are initialized to be 0. We then encode these packets and generate \(k\) redundant packets using RS. Provided that the loss probability computed according to the receiver’s feedback is \(p\). Then

\[(n + k) \times (1 - p) \geq n\]  

(3.1)
Before these packets are delivered, the attribute CSeq of RTP header are used to mark the sequence of packet. After transmission, these packets arrived at the receiver side disorderly. Hence, a jitter buffer is employed at the receiver side to sort these packets into a corrective order. When the jitter buffer is full, we can obtain the total number of packets $M$ sent to receiver by using the CSeq value of last packet to subtract the one of the first packet in jitter buffer. From the size of jitter buffer, we can get the real number $m$ of packets that have been received. So the packet loss rate can be calculated as follows:

$$p = \frac{M - m}{M} \quad (3.2)$$

Based on the packet loss rate, the redundant packet number can be computed as follows:

$$k_i = \frac{n}{1 - p} - n \quad (3.3)$$

That is, the sender should produce at least $np(1-p)$ packets to ensure enough packets arriving at receiver side and being decoded by RS. However, since the networks condition changes incessantly and the packet loss rate also changes all the time. If we use this model to calculate the number of redundant packets, the result will be imprecise. In order to conquer this problem, we propose a Binomial Probability Redundancy Model.

2) Binomial Probability Redundancy Model

Assuming that $n$ is the number of original media data packets, $k_i$ is the number of redundant packets calculated by LPRM, $k_2$ is the real number of redundant packets. As the networks condition varies frequently, a packet may be lost in the networks. We assume that the probability of a packet sent from peer and reach at the receiver side is $p$. And then, the probability that it will be lost is $(1-p)$. As the event of different packets arriving at the destination address is independent, in terms of binomial distribution, the probability of $i$ packets of $(n+k_2)$ arriving is:

$$P_{arr}(i) = C_{n+k_2}^i p^i \times (1 - p)^{n+k_2-i} \quad (3.4)$$

Accordingly, the probability of at least $n$ packets arriving should be:

$$P_{arr} = \sum_{i=n}^{n+k_2} P_{arr}(i) \quad (3.5)$$

Therefore, the general probability of $P_{arr}$ is as follows:

$$P_{arr} = \sum_{i=n}^{n+k_2} C_{n+k_2}^i p^i \times (1 - p)^{n+k_2-i} \quad (3.6)$$

From the equation of $(3.6)$, let $k_2 = k_i$, if we assume $P_{arr}$ is larger than a threshold $\theta$, we can finally compute the minimal number of redundant packets $i$ by using a traversal searching algorithm.

$$P_{arr} \geq \theta \quad (3.7)$$

B. Uni-directional Feedback

The data are sampled from hard device and encapsulated into RTP packets. The sender will send out these packets including redundant packets with continuous increasing CSeq value. The receiver then sorted these packets in the jitter buffer. As the buffer is full, these packets are fetched out for playing. At the same time, the receiver sends back CSeq information as a feedback for sender to adjust RS parameters. We then propose a UDP-based reliable transmission scheme, which employs Binary Exponential Backoff Algorithm (BEBA) [16] to retransmit the loss feedback packets. It mainly contains three types of message: Request, Response and ACK. The message process is shown in Figure 1.

![Figure 1. Processing of Uni-direction Feedback](image-url)

This feedback information contains packet loss rate. If the network is congested, these packets may be lost on the Internet. In our scheme, if the packets are lost, we will trigger a timeout event and retransmit these packets at receiver side. In other words, the sender may obtain many duplicate packets. Since the number of redundant packets is dependent on these feedback packets, if we do not take actions to distinguish the duplicate packets, they will greatly affect the efficiency of FEC RS. Our solution is to append a 4-byte attribute at the signal packet header, named Signal Packet Identifier (SPI). Before the signal packet is sent, SPI is filled with an unsigned integer random number. Obviously, the range of SPI should be $0$ to $2^{32}$. This means that the probability of any pairs of SPIs collide with each other is $1/2^{32}$. The probability is so small that we can consider it as a small probability event. Therefore, two different signal packets are considered to own different SPI values. If the sender receives two packets owning same SPI, one of them should be deemed to be a resend packet and discarded.

IV. MEDIA TRANSMISSION STRATEGY

In this section, we will describe how to transmission the video and audio packets with their redundant packets. We also use an efficient bandwidth control strategy to adjust the quality of video data and the sending rate of sender side:

A. Video Frame Partition

In this paper, we employ H.264 compression standard to compress video data. This standard achieves high compression ratios by exploiting spatial and temporal redundancies in consecutive video frames. Frames contain the whole pictures. Therefore, this type of frame is usually
very large after compression with size of QVGA[19]. Its compressed size is normally from 5 kbyte to 10 kbyte.

Another reason we choose H.264 compression standard is that it contains a number of new features that allow it to compress video much more effectively and an in-loop deblocking filter which helps prevent the blocking artifacts common to other DCT-based image compression techniques, resulting in better visual appearance and compression efficiency.

In our system, the sender produces $k_2$ redundant packets according to BPTM model and the feedback from receiver. Both the redundant packets and raw data are sent to the other peer. While sending data, sender marks the CSeq value of redundant packets to follow the original packets. Just as shown in Figure 2, provided that there are four original packets $(P_1,n_1), (P_2,n_2), (P_3,n_3), (P_4,n_4)$, in which $P$ represents the packet and $n$ represents the CSeq value. By using FEC RS encoder, we can produce the series packets $(R_1,n_3), (R_2,n_6), (R_3,n_1)$, in which $R$ represents the redundant packet and $n$ represents CSeq value.

Rule 1: Each redundant packet are divided into $k$ parts, if the size of packet is not a multiple of $k$, the packet will be swelled with 0 to meet with that prerequisite.

Rule 2: The pieces with different index $i$ from each packet are then picked and encapsulated into a new packet, in which $1 \leq i \leq k$.

Based on the above rules, the $k_2$ redundant packets are regrouped into $k$ new packets. These new packets are then packed into the end of the original data packets. The partition process is shown in Figure 3.

C. Bandwidth Control

To control bandwidth consumption, we propose to adjust the quality of video codec. As is known, the congestion of Internet varies all the time. In the condition of little congestion, the sender only needs to generate a few redundancy packets according our BPRT model. At the same time, we also improve the quality of video codec in order to control the sending rate. Thus, the bandwidth resources are utilized efficiently and good video pictures are displayed at the receiver side.

Provided that $k$ denotes the number of raw audio packets and $k_2$ is the number of redundant packets, then

$$W_{new} = \begin{cases} W_{cur} + k & 0 \leq p \leq \alpha \% \\ W_{cur} & \alpha < p \leq \beta \% \\ \frac{W_{cur}}{2} & \beta < p \leq 100\% \end{cases} \tag{3-8}$$

We use a linear increasing model to control the sending rate. In the equation (3-8), $W_{new}$ represents the new sending rate calculated by BPTM model. When the network is less crowded, the value of $p$ is small. If $p$ is smaller than $\alpha \%$, then we improve the quality of video frame by adding more information to one frame. However, when $p$ is greater than $\beta \%$, sender should reduce the quality of video to avoid the network congestion. The increasing congestion of Internet causes the packet loss rate become larger than before. As a result, we have to

![Figure 2. Video Redundant Processing](Image 270x40 to 350x60)

![Figure 3. Interweaving Recovery Technology](Image 333x484 to 338x499)
generate a large number of redundancy packets to recover the lost ones at the receiver side. This will cause the Internet to be even more crowded. Therefore, we apply BEBA algorithm to control sending rate. Once we detect that the network becomes congestive, the current sending rate \( W_{\text{cur}} \) is divided to create the new sending rate \( W_{\text{new}} \).

V. PERFORMANCE EVALUATION

A. Analytical Evaluation

We analyze the complexity of our model in this section. When sender calculates the arriving probability of continuous \( i \) packets \( \text{Parr}(i) \) by using the equation of (3.4), the complexity of this operation is approximately \( \Theta(n+k^2) \). Hence, when the minimal number of redundant packets is computed using (3.5), the complexity of the expression is nearly \( \Theta((n+k^2)^2) \). In which, the value of \( n \) can be dynamically adjusted. The range of \( n \) is usually \( 3 \leq n \leq 20 \). Therefore, the calculation ability of current mobile terminal is strong enough to meet with this requirement.

B. Identify the Headings

In the experiment, we use two Personal Computers as two peers. The two peers use the real network to communicate with each other. We also deploy ZXDSL 831 in Shenzhen as one of routers to route packets from Shenzhen PC. The following figures show the result of the experiment.

We assume that \( P_{\text{HK}} \) represents one peer running on the PC located in HK and \( P_{\text{SZ}} \) represents the one running at SZ. As the packet loss rate larger than 90% is scarce, it is hard to encounter. Therefore, we deploy a random packet lost strategy on PG to simulate the different network conditions. Because the IP address of \( P_{\text{HK}} \) is a public one, which can be accessed directly from Internet, we firstly start \( P_{\text{HK}} \) listening for connection request, and then start \( P_{\text{SZ}} \) connecting to \( P_{\text{HK}} \).

In order to make a comparison with LPRM under the same network conditions, we respectively run two peers at each side. They are peers of \( P_{\text{HK1}}, P_{\text{HK2}}, P_{\text{SZ1}}, P_{\text{SZ2}} \). \( P_{\text{HK1}} \) and \( P_{\text{SZ1}} \) use LPRM to transmit packets. \( P_{\text{HK2}} \) and \( P_{\text{SZ2}} \) use BPRM. In the experiments, we gather statistical results from two aspects: the number of redundant packets and the arriving probability of packets.

The ITU-T standard of low Bit-rate video codec [16] is employed to encode and decode video data. Besides, H.264 [17] is another ITU-T standard. The reason why we choose H.264 is that the quality of video from H.264 is perfect and the compression algorithm is much more efficient than H.263. Therefore, we use H.264 as the standard codec in our system. Since the size of I frame changes frequently, we only show the results of \( n = 4 \).
In the transmission of audio data, we use Adaptive Multi-rate (AMR) [18] as our audio codec standard. According to AMR standard, the data after encoding is small. Thereby, sending audio in the way of video transmission will waste lots of bandwidth overhead. We compute redundant packets from every 10 audio packets and reconstruct these packets using interleaving technology. We finally piggyback the new packets into original packets. The statistical results are shown in Figure 5.

To simulate the real network conditions, we deploy an active packet loss strategy on PG. From Figure 4(a) and 5(a), we can see that when the packet loss rate is larger than 50%, the number of redundant packets of BPTM is obviously larger than that of LPTM. However, the arriving rate of LPTM decreases gradually. The one of BPTM is still maintained upon 90%. This means that when networks become increasingly congestive, the peer using LPTM may probably not be able to derive the original data from received packets since many packets are lost. The play result will be affected seriously. Especially, when I frame can not be recovered, the video graph displayed on the PC will be filled with large number of Mosaic. Nevertheless, by using BPTM, a significant number of packets will be sent to receiver to guarantee the key frames of video could be derived correctly and played out smoothly.

VI. CONCLUSION

In this paper, we firstly study the linear probability redundancy model on the aspect of bandwidth utilization. Through the experiments in the real networks, we find that LPRM works poor when the networks become congestive and the loss of packets become common. To fill this deficit, we propose a binary probability redundancy model. In our model, the networks congestion factor is fully taken into consideration. The sender produces enough redundant packets according to the networks congestion to guarantee the receiver can derive the original data from the received packets. Even though BPTM can work efficiently, the large number of audio redundant packets still will worsen the network. Therefore, we will focus on how to cut down audio raw data to decrease the number of redundant packets in our future works.

VII. REFERENCES