Adaptive FEC-Based Packet Loss Resilience Scheme for Supporting Voice Communication over Ad Hoc Wireless Networks

Venkat Raju Gandikota, Bheemarjuna Reddy Tamma, *Member*, *IEEE*, and C. Siva Ram Murthy, *Senior Member*, *IEEE*

Abstract—Providing real-time voice support over multihop ad hoc wireless networks (AWNs) is a challenging task. The standard retransmission-based strategies proposed in the literature are poorly matched to voice applications because of timeliness and large overheads involved in transmitting small-sized voice packets. To make a voice application feasible over AWNs, the *perceived voice quality* must be improved while not significantly increasing the packet overhead. We suggest packet-level media-dependent adaptive forward error correction (FEC) at the application layer in tandem with multipath transport for improving the voice quality. Since adaptive FEC masks packet losses in the network, at the medium access control (MAC) layer, we avoid retransmissions (hence, no acknowledgments) in order to reduce the *control overhead* and *end-to-end delay*. Further, we exploit the combined strengths of layered coding and multiple description (MD) coding for supporting error-resilient voice communication in AWNs. We propose an efficient packetization scheme in which the important substream of the voice stream is protected adaptively with FEC depending on the loss rate present in the network and is transmitted over two maximally node-disjoint paths. The less important substream of the voice stream is encoded into two descriptions, which are then transmitted over two maximally node-disjoint paths. The performance of our scheme (packet-level media-dependent adaptive FEC scheme) is evaluated in terms of two parameters: *residual packet loss rate* ($RPLR$, *packet loss rate after FEC recovery*) and *average burst length* (ABL, *average length of consecutive packet losses after FEC recovery*) of voice data after FEC recovery. The sets of equations leading to the analytical formulation of both $RPLR$ and ABL are first given for a renewal error process. The values of both these parameters depend on $FEC\text{-}Offset$ (r, the distance between original voice frame and piggybacked redundant voice frame) and loss rate present in the network. Then, these parameters are computed for a Gilbert-Elliott (GE) two-state Markov error model and compared with experimental data. Our scheme adaptively selects the $FEC\text{-}Offset$ (it chooses r that minimizes $RPLR$ and ABL as much as possible) based on the loss rate feedback obtained from the destination. The proposed scheme achieves significant gains in terms of reduced *frame loss rate* (FLR), reduced *control overhead*, and minimum *end-to-end delay* and almost doubles the *perceived voice quality* compared to the existing approaches.

Index Terms—Ad hoc networks, voice frame, layered coding, multiple description coding, forward error correction, packetization scheme, voice quality, multipath transport, multimedia.

1 Introduction

Ad hoc wireless networks (AWNs) are formed by a set of mobile nodes that communicate with each other over a wireless channel without the help of any preexisting infrastructure. Nodes cooperate to forward packets for effecting communication between any two nodes that are not directly within the wireless transmission range of one another. Since these networks can be deployed rapidly and flexibly, they are attractive for numerous potential applications, ranging from emergency and rescue operations to real-time multimedia communications for disaster areas. Real-time multimedia applications can tolerate packet losses to some extent (up to 5 percent) but are highly delay sensitive (typically for interactive voice communication, the *end-to-end delay* should be less than 200 ms [1]). In this paper, we focus mainly on providing voice support over AWNs, because it (voice over an AWN) is potentially a key application in many current and proposed scenarios. When transmitting voice data, continuous delivery with limited packet loss rate is of primary importance than trying to achieve zero packet loss rate by employing retransmission-based strategies. For a packet to be useful at its destination, the destination should receive the packet before its playout time. Playout time is defined as the time instant at which the packet should be played out at the destination. Thus, each packet has to reach the destination within the specified deadline (its playout time), after which it becomes useless.

The unique characteristics of the voice communication, such as small payload size (typically 20 bytes) and timely arrival of the packets at the destination, make it very challenging to deploy over AWNs. In wireless networks, in addition to packetization overheads (i.e., header information of the medium access control (MAC) and other higher layers as in wire-line networks), each packet should also include a preamble for synchronization, which typically occurs on a packet-by-packet basis. Also, after receipt of

- *V.R. Gandikota is with Yahoo! Software Development India Pvt. Ltd., Bangalore 560001, India. E-mail: gvrajuindia@yahoo.co.in.*
- *B.R. Tamma is with CalIT2, Jacobs School of Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093. E-mail: btamma@ucsd.edu.*
- *C.S.R. Murthy is with the Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai 600036, India. E-mail: murthy@iitm.ac.in.*

Manuscript received 10 Oct. 2007; revised 11 Feb. 2008; accepted 20 Feb. 2008; published online 7 Mar. 2008.
For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-2007-10-0306.
Digital Object Identifier no. 10.1109/TMC.2008.42.

each packet, the receiver needs to send an acknowledgment (ACK) to indicate a successful reception, which further increases the *control overhead*. Thus, the performance of such networks is poor for small data payloads common in voice communication. In a single-hop network, every node is within the radio range of every other node, and hence, the *end-to-end delay* is not too large. However, in the multihop scenario, the source and destination may be several hops away and packets need to be forwarded by the intermediate nodes, causing delays at each hop. As a result, the *end-to-end delay* can become quite large, especially due to packets of a multihop flow contending with each other for the shared channel at successive hops, thereby making voice communication very challenging. The quality of voice call degrades as packet loss and delay increase. Extreme losses may render the speech unintelligible. Likewise, as the *end-to-end delay* increases, the interaction between call participants becomes more difficult to establish.

In this paper, we propose a scheme which effectively combines adaptive forward error correction (FEC), layered coding, and multiple description (MD) coding to achieve significant performance gains in terms of *perceived voice quality*, *frame loss rate* (FLR), and *end-to-end delay*. Given a voice stream, using a layered coder (LC), it can be encoded into two substreams, viz., important and less important substreams [2]. The important substream and less important substream are called base layer (BL) substream and enhancement layer (EL) substream, respectively. In the proposed scheme, the BL substream is protected with an adaptive FEC mechanism and is transmitted over two maximally node-disjoint paths in order to increase the success probability of packet reception. The EL substream is further encoded into two descriptions using MD coder, and each description is sent along with the BL substream. Each voice packet contains the adaptive FEC-protected BL and one of the two EL descriptions. As we explain later (in Section 3), due to this effective packetization, better *perceived voice quality* can be achieved as the scheme is more immune to packet losses in the network. The rest of this paper is organized as follows: Section 2 briefly discusses the related work. Section 3 provides a detailed description of the proposed packetization scheme. Sections 4 and 5 discuss the analytical framework and simulation results, respectively. Finally, Section 6 contains concluding remarks.

2 Related Work

Some of the earlier works [3], [4], [5] have addressed the issue of supporting voice communication over AWNs. These works focus on optimizing packet length, employing forward error control within a packet, reservation policies, bandwidth reuse technique, and retransmission strategies. In [4], a single description coded video transmission system is modeled and analyzed using a combination of packet-level FEC and path diversity. The authors of [5] proposed a proxy-based solution for enhancing the performance of UDP/IP traffic over 802.11b-based wireless networks. In [6], a low-delay interleaving and conditional retransmission scheme was proposed to improve the video quality for wireless video. The impact of using media-dependent FEC in VoIP flows over the Internet was discussed in [7]. The authors of [8] deal with energy consumption of FEC on wireless devices, whereas the authors of [9] describe a hybrid simulation tool for the evaluation of voice transmission through a large wireless network with different network parameters such as mobility and traffic congestion.

All the above works do not explicitly concentrate on reducing the large overheads incurred by small-sized voice packets and do not exploit the combined strength of FEC, layered coding, and MD coding for supporting voice communication over AWNs. Even though there exist several multipath routing protocols for AWNs, the focus has mainly been on the support of delay-insensitive data applications, rather than on improving end-to-end performance for real-time traffic.

A potentially promising approach to reduce the voice packet loss rate is to establish multiple paths between the source and destination of a session and to use speech coding schemes that take advantage of the existence of multiple paths. One such coding scheme is MD coding [10], in which a voice stream is encoded into multiple independent substreams (descriptions). These descriptions can be decoded independently to produce a voice stream of basic quality. MD coding does not require prioritized transmission as all descriptions have equal importance. When more descriptions are received, the decoder can gradually increase the quality. Since the probability of losing all descriptions is relatively low, it performs better than LC at higher packet loss rates. MD coding has a rich history [11], [12], [13] and has been studied extensively in the literature. However, there is no prior work that integrates design concepts across multiple layers to provide an integrated system suitable for supporting real-time voice applications over AWNs.

In [3], a combination of interpacket redundancy, MD coding, and path diversity was used to provide speech support over AWNs. However, the authors modified the IEEE 802.11 MAC protocol to provide interpacket redundancy. By applying MD coding on the whole voice stream, they obtained two descriptions which are then, respectively, transmitted over two paths set up by the routing module. Our work mainly differs with [3], in terms of providing an efficient packetization scheme wherein we use adaptive FEC at the BL and MD coding at the EL. The idea of protecting the BL substream strongly when the loss rate in the network is high by using adaptive FEC (alternatively the BL substream is less protected when the loss rate is low) and applying MD coding to the EL substream to improve the perceived voice quality have not been exploited particularly for transmitting voice over AWNs in earlier attempts. Preliminary results show that by exploiting the strengths of FEC, LC, and MD coding, the mechanisms proposed in our scheme can indeed provide better voice quality and reduced packet loss rate than other related approaches [2], [3] for real-time voice.

3 An Efficient Packetization Scheme

During voice communication, if the number of lost voice packets is higher than that tolerated by the listener, then either an error control or loss recovery mechanism is required. Typical mechanisms fall in one of the two classes, viz., closed-loop mechanisms and open-loop mechanisms [14]. Automatic Repeat Request (ARQ) mechanisms are closed-loop mechanisms where the source retransmits lost packets as reported by the destination. ARQ mechanisms

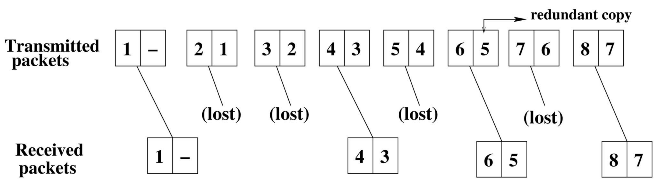

Fig. 1. Typical media-dependent FEC mechanism.

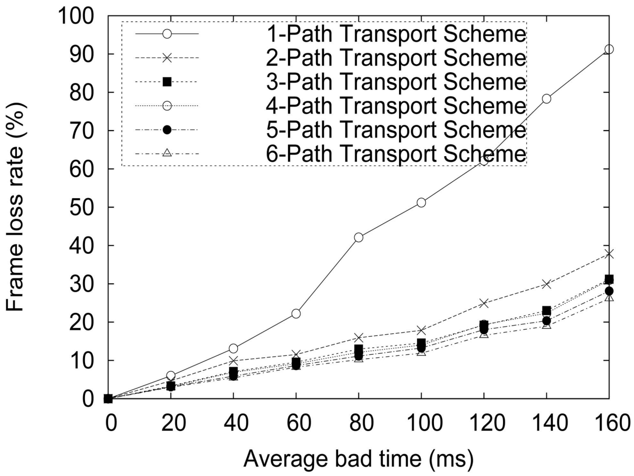

Fig. 2. FLR versus ABT.

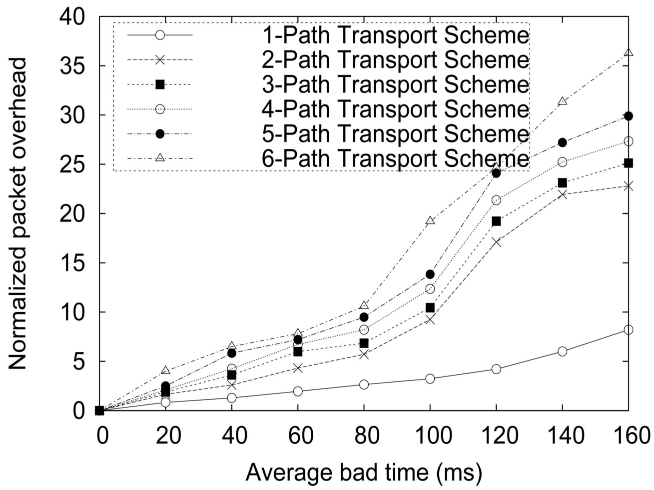

Fig. 3. NPO versus ABT.

are typically not acceptable for interactive speech communication because they increase the *end-to-end delay*, and thus, the packets might miss the deadline. FEC mechanisms are open-loop mechanisms, where redundant data is transmitted along with the original data so that (at least some of) the lost original data can be recovered from the redundant data. FEC mechanisms can be further classified into two categories: media independent and media specific. In media-independent FEC-based methods, the redundant information can be sent in the form of parity packets [15]. The parity packets are generated from the original packets using a mathematical function, and thus, the redundant information is independent of the captured voice data. However, media-independent FEC schemes are not well suited for interactive voice because they require that data to be broken up into blocks, which in practice would be of large size. Thus, the use of such schemes would add a nonnegligible *block delay* (delay incurred while waiting for packets that belong to a block) to the *end-to-end delay*, thereby affecting the interaction between the participants. Media-specific schemes piggyback information about the voice packet(s) that correspond to present period with later packets, as shown in Fig. 1. If a packet $n-i$ that has a redundant encoding copy in packet n is lost, then the application has to wait for packet n to recover from the loss. Thus, the application can recover a lost packet with i packets worth of delay. We use media-specific FEC for providing protection to the voice stream in our scheme.

3.1 Motivation

Our proposed scheme is based on the following observation. In layered encoded voice communication, the BL substream has a special significance over EL substream. If some or all the bits of EL substream are corrupted or lost and if only the BL substream is available at the destination, then it can still decode the BL substream to obtain degraded but acceptable quality (i.e., base quality). Therefore, if the BL substream is protected adaptively based on the packet loss rate in the network, then at least the base quality can be guaranteed for the ongoing session. However, the amount of protection should always be kept low in order to reduce the overhead. To increase the *perceived voice quality* while minimizing the redundancy (packet overhead), the EL substream is encoded into two descriptions using MD coding (we explain this in Section 3.2) and transmitted along two maximally node-disjoint paths.

In the proposed packetization scheme, we suggested to use multipath transport (to sustain packet losses in the network due to path breaks and lossy links) at the network layer. In particular, we used two maximally node-disjoint paths between source and destination nodes instead of three or four paths. As mentioned in [16], a maximum of 50 percent improvement in throughput can be achieved in a multihop wireless network if we use two independent paths between source and destination, and transmit data in a round-robin fashion. If we set up more than two paths between source and destination nodes, then it will increase the overall network traffic, leading to collisions and congestion, and, thus, affecting the performance of the network.

To demonstrate the effectiveness of two-path transport scheme, we ran experiments to measure FLR, normalized packet overhead (NPO), voice quality in terms of Perceptual Evaluation Speech Quality Mean Opinion Score (PESQ-MOS), and number of collisions per node for varying *average bad time* (ABT) under 1-path, 2-path, 3-path, 4-path, 5-path, and 6-path scenarios. For these experiments, we transmitted the full voice packets along all paths in a redundant fashion. Figs. 2, 3, 4, and 5, respectively, show the changes in FLR, NPO, PESQ-MOS, and a number of collisions with respect to *ABT*. FLR is calculated as $[1 - \text{Frame Delivery Ratio (FDR)}]$, where FDR is defined as the ratio of the number of voice frames received within the deadline by the application layer of the destination to the number of voice frames sent by the application layer at the source node. NPO is defined as the ratio of the total number of packets (control and data) exchanged over the total number of data packets received in the network. The PESQ-MOS score [between -0.5 (worst) to 4.5 (best)] is evaluated using the ITU-T perceptual measurement algorithm. A detailed explanation on PESQ-MOS is

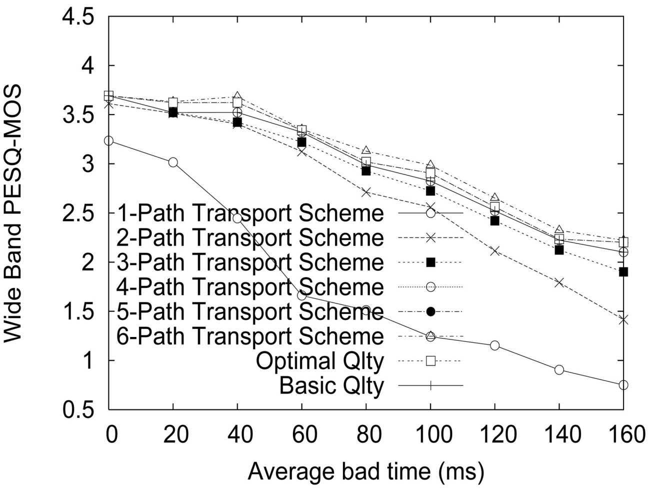

Fig. 4. PESQ-MOS versus ABT.

given in Section 5.3. At each node, the total number of collisions occurred is measured and is used for computing the average number of collisions per node in the network. This metric does not include those packets that are dropped due to bit errors. We used the NS-2 network simulator [17] to simulate the above scenarios. The simulation parameters are specified in Table 1. Mobility is simulated according to the Random Way Point model [18]. According to this model, the nodes remain stationary for a certain pause time after which they move in a randomly chosen direction with a random velocity chosen uniformly between a specified minimum and maximum velocities. For all cases of mobility in the network, we set the pause time to 0 and set the minimum and maximum speeds to the same value to ensure that the nodes move at a constant speed. We use Adaptive Multi-Rate Wide Band (AMR-WB) speech codec with 12.65 Kbps bit rate with a sample size of 253 bits ($\approx$ 32 bytes) for sending voice packets from the source to the destination. Voice packets are generated at the source following constant bit rate (CBR) pattern with an interframe time of 20 ms (i.e., 50 voice packets per second). We have 10 voice flows in the network that start at random times, and in each voice flow, 15,000 voice packets are sent from the source to the destination. The total

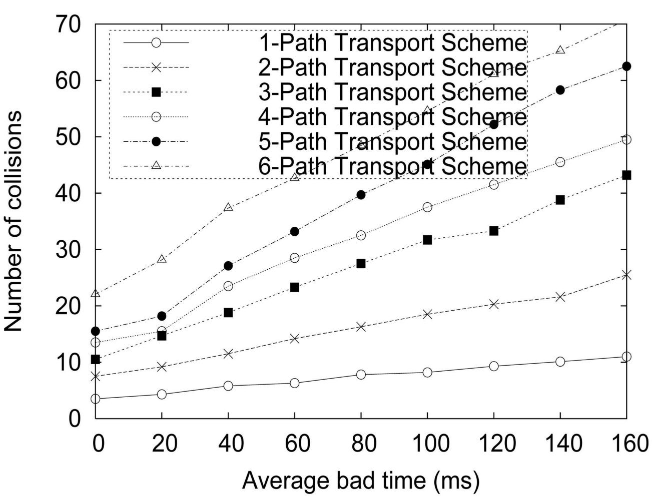

Fig. 5. Number of collisions versus ABT.

TABLE 1
Simulation Parameters

Parameter	Value
Terrain Area	1,000 *m* x 1,000 *m*
Transmission Range	250 *m*
Channel Capacity	11 Mbps
No. of Nodes	75
Mobility Model	*Random Way Point* (RWP)
Simulation Duration	1,000 *s*
No. of Frames per Voice Flow	15,000
No. of Voice Flows	10
MAC Protocol	802.11 DCF
Inter Frame Time	20 *ms*
Traffic Type	Constant Bit Rate (CBR)
Voice Frame Size	253 bits

simulation duration is fixed at 1,000 seconds. The same kind of simulation setup was used in [19] and the authors showed that the performance obtained using a simulation duration of 1,000 seconds is similar to that of 4,000 seconds. The RTS-CTS mechanism is disabled as the voice packet payload size is relatively very small ($\approx$ 32 bytes). We modified the Dynamic Source Routing (DSR) [20] protocol to obtain multiple maximally node-disjoint routes from the source to the destination. In all our simulation experiments, we used the Capture Threshold Model (CTM), the default interference model in the NS-2 simulator. This model is based on "capture." Rather than using a fixed interference range, or assuming that the nodes to be colocated, collisions are determined by comparing the desired signal strength at the receiver, with the level of interference. However, the signal strength of the packet being received is compared with the interfering power from a single node at a time, rather than using the Signal-to-Interference-to-Noise Ratio (SINR) with cumulative interferences, to determine successful reception. In spite of its simplistic interference model, the NS-2 is being widely used to evaluate new proposals for routing and MAC protocols. Since we used the CTM for evaluating all the schemes, the CTM overestimates the optimal performance for all the schemes consistently, and hence, conclusions drawn in this work are still valid. All the results that we have shown contain data points averaged over 10 flows. In all our simulation experiments, we considered that the source and destination are several hops away from each other. Simulation runs are carried out for 30 seeds, and all the results conform to 95 percent confidence levels. We used the modified *Gilbert-Elliott* (GE) two-state Markov error model [3] available in the NS-2 for each link to simulate bursty packet losses in the network. When the state of a path is bad, all the packets are lost; when the state is good, the packets can still be lost due to collision or noise. The average dwell time in the good state is 1,000 ms. The average dwell time in the bad state (*ABT*, the average time spent in the bad state) is varied to simulate bursty packet losses in the network. Unless otherwise stated, the same simulation setup is used in all our subsequent simulation experiments.

Table 2 quantitatively compares various schemes with respect to 1-path and 2-path transport schemes. As observed from the table, an increase in the number of paths results in improvement in FLR and voice quality with respect to 1-path scheme. However, a significant improvement is achieved when we move from 1-path to 2-path

TABLE 2
Quantitative Comparison of Various Transport Schemes

Schemes Compared	FLR (% Reduction)	NPO (% Increment)	PESQ-MOS (% Increment)	No. of Collisions (% Increment)
2-path vs 1-path	50.48	129.44	83.69	116.71
3-path vs 1-path	52.60	213.27	87.64	256.66
4-path vs 1-path	54.16	232.26	92.13	324.93
5-path vs 1-path	57.79	277.41	96.65	414.00
6-path vs 1-path	59.92	439.41	97.65	714.00
3-path vs 2-path	4.20	64.76	4.71	119.9
4-path vs 2-path	7.29	79.43	10.08	178.39
5-path vs 2-path	14.48	114.32	15.48	254.70
6-path vs 2-path	18.71	239.47	16.68	511.78

transport scheme. Further, the number of collisions increases drastically as we increase the number of paths in the transmission. Though 3-path (also 4-path, 5-path, and 6-path) scheme slightly improves FLR and voice quality with respect to 2-path scheme, it results in a significant increase in NPO and a number of collisions. Thus, the usage of 2-path transport scheme is more appropriate for voice transport in AWNs.

3.2 Combining the Strengths of Layered Coding and MD Coding

By exploiting the strengths of both LC and MD coding, we propose an effective packetization scheme at the application level as follows:

1. BL substream is protected using traditional adaptive FEC and transmitted over two maximally node-disjoint paths. Depending on the packet loss rate in the network, the *FEC-Offset* (see Fig. 6) is varied, and thus, the BL is strongly protected.

2. EL substream is encoded into two descriptions to take advantage of multipath transmission and to reduce the packet overhead. Each path carries one of the two descriptions of the original EL substream. When both the descriptions are received at the destination, the destination voice application can retrieve the original EL substream from the two descriptions. However, if only one of them is received, then the received description contributes in improving the quality of the BL substream. Each transmitted packet contains the FEC-protected BL and one of the two descriptions of the EL. Thus, with high probability, both the BL substream and one of the two descriptions of the EL substream are received (even in the presence of packet losses in the network), and hence, the quality of the received stream is better than the base quality. To illustrate the above techniques, we consider the AMR-WB speech codec [21] with 12.65 Kbps bit rate. The AMR-WB speech codec has been originally developed by the Third Generation Partnership Project (3GPP) to be used in GSM and 3G cellular systems. The coding scheme is the Algebraic Code Excited Linear Prediction. Voice activity detection, comfort noise generation, source-controlled rate operation, and error concealment of lost frames are also provided in the specifications. The multirate codec has eight encoding modes ranging from 6.6 to 23.85 Kbps. The codec is adaptive in the sense that it can switch its bit rate every 20 ms of speech data depending upon channel and network conditions. At the output of the encoder, bits are ordered according to their subjective importance and further divided into three classes with decreasing perceptual importance of Class A, Class B, and Class C. The AMR-WB (for a 12.65-Kbps rate) speech codec produces a voice packet of size 253 bits, as shown in Fig. 7a. The 253 bits of the AMR-WB voice packet can be classified into 72 important (BL or Class A) bits and 181 less important (EL or Class B) bits, as shown in Fig. 7b. There are no Class C bits in this mode. The packet formats for paths 1 and 2 are shown in Figs. 7c and 7d, respectively. For the sake of simplicity, we employ even-odd decomposition method to get two descriptions from the EL bits, as shown in MDC-1 and MDC-2 parts of Figs. 7c and 7d, respectively.

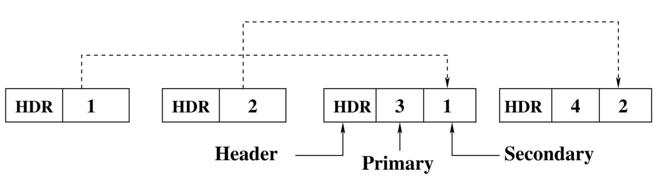

Fig. 6. FEC mechanism with offset 2.

3.3 Working Mechanism of Packetization Scheme

As shown in Fig. 8, AMR-WB Layered Codec takes a raw voice frame as the input and produces the BL substream and one EL substream. The EL substream is encoded into two descriptions by the EL Multiple Description Codec (EL-MD Codec). The source node then creates two voice packet entities by using the BL substream and the corresponding two descriptions (Figs. 7c and 7d). Both voice packet entities contain the same sequence number. Each of these voice packet entities is then encapsulated into a UDP packet and then sent down to the network layer. Network layer then tries to transmit them on two maximally node-disjoint paths. If only single path exists then both the voice packet entities that correspond to the same voice frame will be routed along the same path. At the MAC layer, we use 0-retransmissions (we will explain the reason for this in the next section). At the destination side before accepting a voice packet, its timestamp is checked to see whether it is received before the deadline or not. In case it has missed its deadline, the packet is dropped.

If both voice packet entities (that contain the same sequence number and, thus, correspond to a single voice frame) are received within the deadline by the destination voice application, the EL-MD Codec then recombines the two EL descriptions of these two voice packet entities to get the original EL packet. If it receives only one of the voice packet entities, then the bits corresponding to the other description are made zeros, before the EL packet is given to AMR-WB Layered Codec. The AMR-WB Layered Codec

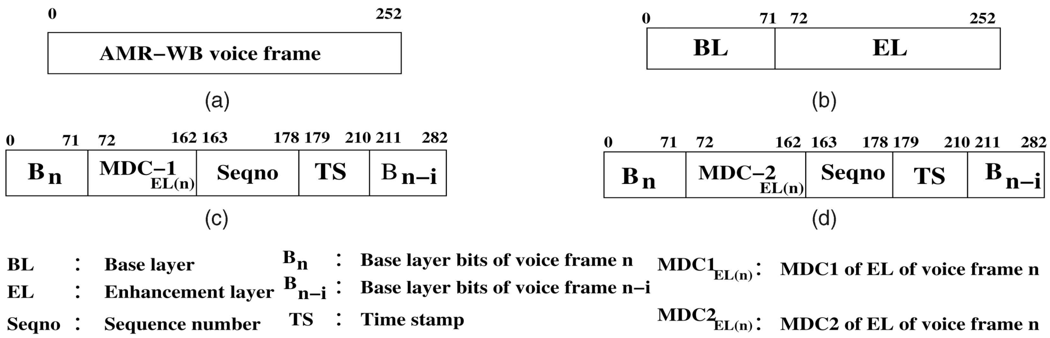

Fig. 7. Packet formats. (a) AMR-WB voice packet. (b) Layer wise bit classification. (c) Packet format for path 1. (d) Packet format for path 2.

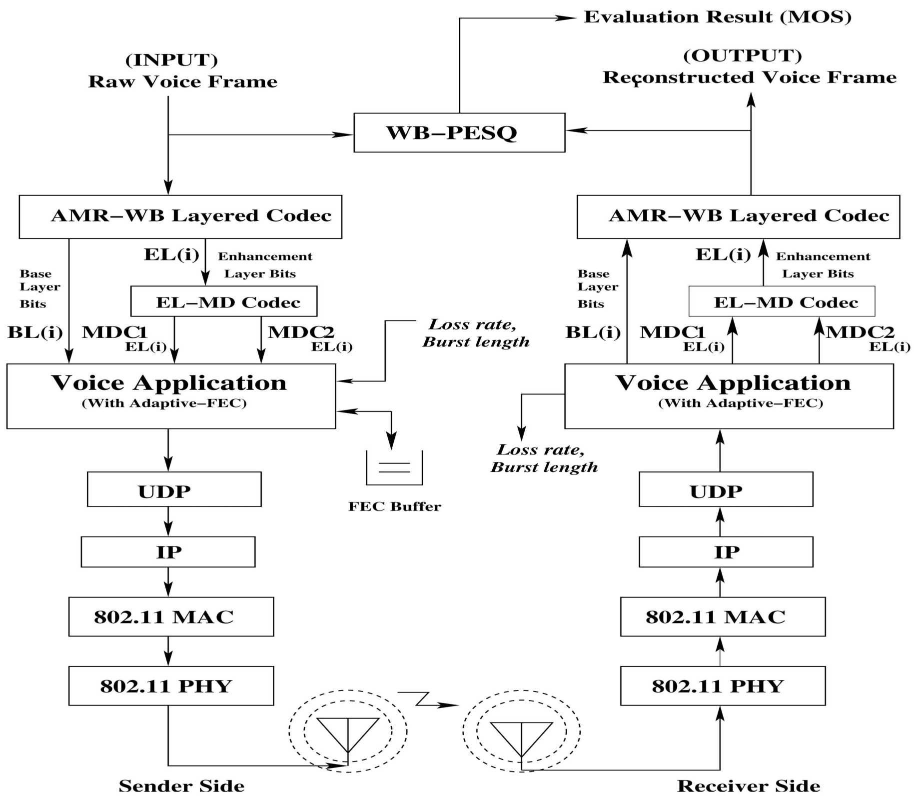

Fig. 8. Layer wise description.

combines the BL and EL packets to obtain the original (or degraded) voice frame. On the other hand, if both the voice packet entities are lost in the transmission, then the voice application uses FEC to recover from the loss of voice frame.

For the redundant information to be most effective, the *FEC-Offset* should be varied dynamically based on the actual loss rate in the network. The voice application at the source node should estimate the target perceived loss rate (i.e., loss rate after FEC decoding) at the destination and it should choose the optimal *FEC-Offset* that will yield the target perceived loss rate closest to the minimum. The voice application module at the source holds *FEC-Offset* parameter and *FEC Buffer* to serve the purpose of adaptive FEC mechanism. The *FEC Buffer* temporarily holds the recently transmitted voice frames so that these frames can be piggybacked to the ongoing voice packets as FEC data. The *FEC-Offset* is used for piggybacking FEC data adaptively along with the original voice frame. The voice application module finds the best possible *FEC-Offset* based on the network loss parameters p and q received from the destination node. These parameters p and q denote the transition probabilities between states 0 (corresponds to correct reception) and 1 (corresponds to loss of a packet), respectively, in a GE model. The GE model is a two-state Markovian error model [22] with geometrically distributed residence times and is shown in Fig. 9. The parameter p is defined as the probability that the next packet is lost, provided the previous one has arrived, and parameter q is defined as the probability that the next packet is received, provided the previous one has lost (the following section describes how p and q values are calculated from the packet history and also illustrates the adaptive FEC mechanism). We propose the following schemes for providing the feedback information (network loss parameters, p and q) to the source node:

- *Scheme for one-way voice communication*. In this scenario, the destination node periodically sends

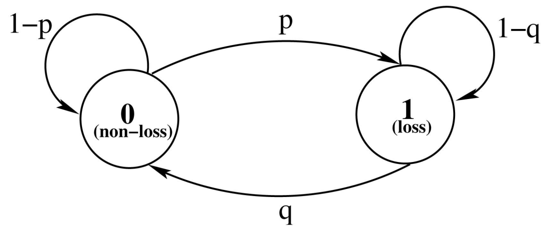

Fig. 9. GE model.

an explicit packet containing feedback parameters to the source to reflect the network state (due to the absence of reverse direction traffic).

- *Scheme for interactive voice communication*. In this scenario, the feedback parameters are piggybacked to the incoming voice packets, and thus, the overhead is reduced compared to the previous scheme. However, to tackle the longer periods of silence in the reverse direction, we could use a timer called MAX_SILENCE_TIMER. If this timer expires, then the destination node sends an explicit packet containing the feedback parameters to the source node without waiting for an outgoing voice packet to piggyback these parameters.

3.4 Adaptive FEC Mechanism of Packetization Scheme

In this section, we describe how the source adaptively adjusts the $FEC\text{-}Offset$, r, for reducing the packet loss rate after FEC recovery and the packet overhead. The destination node maintains a packet count history in order to measure the packet loss rate (in terms of p, q) in the network. The packet count history contains the sequence numbers of successfully received voice frames over a time window. On receiving a packet along one of the two paths, the destination node updates the packet count history with the corresponding packet's sequence number. Using the packet count history, the destination node calculates the values of parameters p and q as follows: Let m_i, $i = 1, 2, 3, 4, 5, \cdots, n-1$ denote the number of loss bursts having length i, where $n-1$ is the length of the longest loss burst. Let m_0 denote the number of delivered packets. The parameters p and q are calculated as

$$p = \frac{\left(\sum_{i=1}^{n-1} m_i\right)}{m_0} \quad \text{and} \tag{1}$$

$$q = 1 - \frac{\left(\sum_{i=2}^{n-1} m_i \times (i-1)\right)}{\left(\sum_{i=1}^{n-1} m_i \times i\right)} = \frac{\left(\sum_{i=1}^{n-1} m_i\right)}{\left(\sum_{i=1}^{n-1} m_i \times i\right)}. \tag{2}$$

The destination node periodically calculates these parameters and feeds them back to the source. The media-dependent FEC performance is dependent on the burstiness of the underlying packet-loss process. Generally, the lesser bursty the packet losses are, the better the performance achieved by media-dependent FEC. In order to achieve the best perceived quality, the FEC mechanism must be highly adaptive. The source node cannot blindly increase $FEC\text{-}Offset$, r for achieving better *perceived voice quality* because of the following reasons. First, the higher the $FEC\text{-}Offset$ value, the longer the waiting time that the destination should wait for recovering the lost voice frame. Second, there may be a situation that the *Residual Packet Loss Rate* ($RPLR$, packet loss rate after FEC recovery) is less for a particular $FEC\text{-}Offset$ value but the *Average Burst Length* (ABL, average length of consecutive packet losses after FEC recovery) may be higher (as shown in Figs. 19 and 20, refer to Section 4.5 for detailed explanation).

Hence, while selecting the optimal $FEC\text{-}Offset$, importance should be given to both $RPLR$ and ABL as the packet-loss recovery mechanism at the destination node cannot sustain high burstiness and may not benefit from FEC mechanism. We observe that selecting the $FEC\text{-}Offset$, r that has minimum ABL if the target perceived loss rate of the destination, $RPLR(p, q, r)$ (estimated at the source), is $\leq$ 5 percent (where 5 percent is the maximum tolerable voice packet loss rate) and for all other cases choosing r that has the minimum $RPLR$, provides the best possible perceived voice quality. The selection criteria of $FEC\text{-}Offset$ is summarized in Algorithm 1. Section 4 deals with how we estimate the $RPLR$ and ABL given $FEC\text{-}Offset$ and packet loss parameters p and q.

Algorithm 1. Calculate $FEC\text{-}Offset$ value given (p, q)
Input: Network Loss Parameters p and q
Output: $FEC\text{-}Offset$, r
Begin
  $Offset \leftarrow MAX\_OFFSET$ /* MAX_OFFSET refers to maximum $FEC\text{-}Offset$ value */
  /* Check whether there exists an $FEC\text{-}Offset$ such that $RPLR \leq 5$ percent */
  for $r \leftarrow 1$ to MAX_OFFSET **do**
    /* Calculate the *perceived loss rate*, $RPLR$ and *avg. burst length*, ABL that are seen by the dest. */
    $RPLR[r] \leftarrow RPLR(p, q, r)$ /* refer (6), Section 4 */
    $ABL[r] \leftarrow ABL(p, q, r)$ /* refer (7), Section 4 */
    if $\text{RPLR}[r] \leq 5$ percent **then**
      $FLAG \leftarrow \text{TRUE};$
      $Burst\_Size \leftarrow ABL[r]$
    end if
  end for
  /* Find the $FEC\text{-}Offset$ that has minimum ABL if $RPLR \leq 5$ percent */
  if $FLAG = \text{TRUE}$ **then**
    for $r \leftarrow 1$ to MAX_OFFSET **do**
      if $RPLR[r] \leq 5$ percent and $ABL[r] \leq Burst\_Size$ **then**
        $Offset \leftarrow r$ /* save the $Offset$ value */
        $Burst\_Size \leftarrow ABL[r]$
      end if
    end for
  end if
  return $Offset$
End

3.5 Impact of MAC-Retransmissions on the Network Traffic

An important point to be noted here is that the usage of multipath transport almost doubles the network traffic, even though it provides benefits such as better error resilience, failure recovery (using path redundancy), and

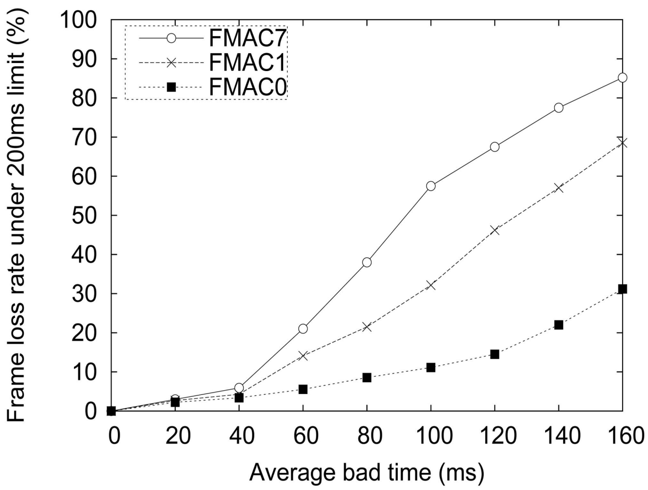

Fig. 10. FLR versus ABT.

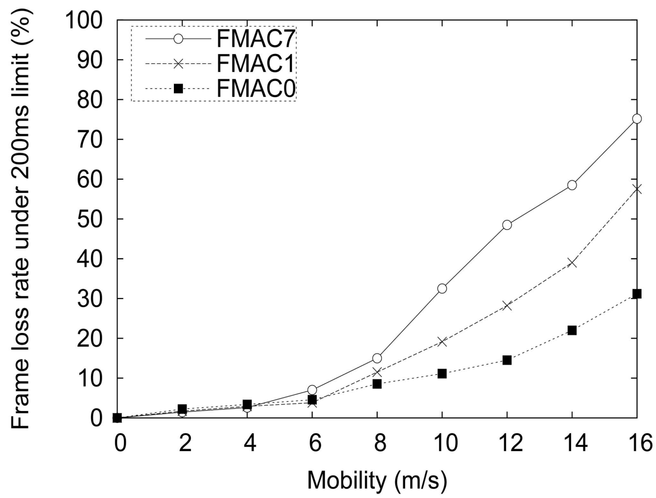

Fig. 11. FLR versus mobility.

reducing correlated losses along a given path. At the MAC layer, if the data packet is successfully received, then the receiver sends an ACK to the sender node. However, if an ACK is not received, then the data packet is assumed to be lost and a retransmission is scheduled by the sender node. In 802.11 MAC standard, up to seven retransmissions are allowed before a packet is dropped.

Since in our scheme we employ multipath transport at the network layer and FEC at the application layer, it conceals packet losses in the network. Therefore, MAC retransmissions are redundant and they unnecessarily increase the network traffic. In a single-hop wireless network, the overhead for transmission of ACK frame is minimal and the contribution to end-to-end delay would not be particularly significant for somewhat infrequent single retransmission attempts. However, in a multihop wireless network, MAC-level ACKs and retransmissions may increase the network traffic and end-to-end delay. Therefore, in our scheme, we disable the ACK mechanism at the MAC layer and set the number of MAC retransmissions to 0 (so that the lost packets will not be retransmitted). This will reduce the unnecessary ACK packet overhead and helps in decreasing the network traffic. If there are any lost packets in the voice stream, then the receiver will try to conceal the packet loss using the FEC mechanism.

In the following, we describe the gains that can be obtained by eliminating ACKs at the MAC layer and employing FEC at the application layer. We explore three schemes to find the effectiveness of FEC at the application layer and the number of retransmissions at the MAC layer. We used the NS-2 network simulator [17], and the simulation parameters are specified in Table 1. For all the simulations, we set $delay_{threshold} = 200$ ms. Refer to Section 3.1 for complete simulation setup details. In the following three schemes, FEC with offset 1 (packet $n-1$ is piggybacked along with packet n) is applied at the application layer. For all the schemes, a voice packet is considered to be lost if redundant copies, both original and reconstructed (after FEC recovery), of that packet are lost. While varying mobility, the ABT is kept constant at 30 ms:

- **FMAC7**. In this scheme, seven MAC-level retransmissions are used at the MAC layer.
- **FMAC1**. In this scheme, one MAC-level retransmission is used at the MAC layer.
- **FMAC0**. In this scheme, MAC-level retransmissions are disabled (and, thus, no ACKs) at the MAC layer.

Fig. 10 shows the *FLR* for varying ABT (the average time spent in the bad state of GE model, see Section 3.1). Since both FMAC1 and FMAC7 schemes employ retransmissions to recover from transmission failures, under high loads they increase network traffic, thereby increasing both *end-to-end delay* and *FLR*. However, FMAC0 scheme shows that by just applying FEC at the application layer and going for 0-retries (no ACKs) at the MAC layer, better *FLR* and *end-to-end delay* can be achieved. Figs. 11 and 12, respectively, show the *FLR* and *end-to-end delay* for varying mobility. As the mobility increases, there are more path breaks which cause burst losses in the network. For both FMAC1 and FMAC7 schemes, the advantage of adding FEC at the application layer is nullified, due to an increase in the network traffic causing more collisions and increasing average burst length. Thus, FMAC0 performs better than both FMAC1 and FMAC7 schemes. Hence, we employ FMAC0 mechanism in our scheme for achieving the best possible voice quality.

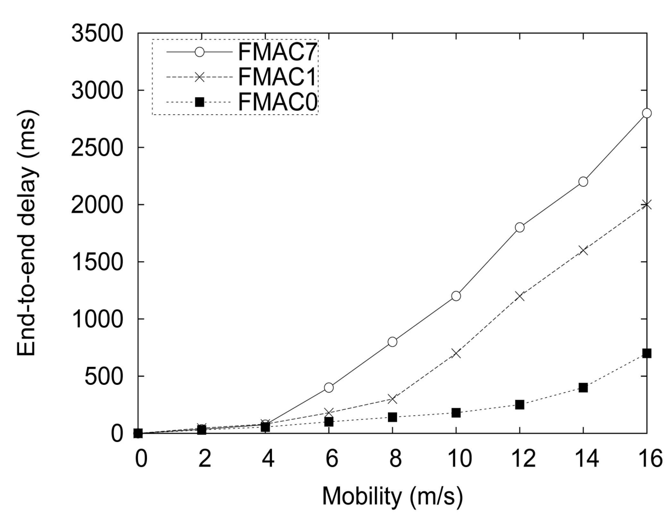

Fig. 12. End-to-end delay versus mobility.

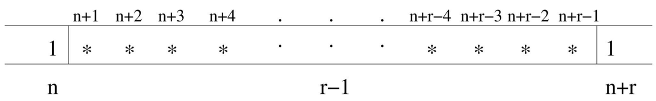

Fig. 13. Calculating residual packet loss rate.

4 Theoretical Analysis

In this section, we derive analytical expressions for the *average burst length* and the *average packet loss rate* by considering the renewal error process. We first provide a brief introduction to the renewal error process and then proceed to analyze the loss pattern in the case of a typical media-dependent FEC. We assume that the loss process can be modeled with a renewal error process. In other words, the lengths of consecutive inter-error intervals (also called gaps) are assumed to be independently and identically distributed. For the sake of clarity, assume that every packet is assigned a binary value 0 or 1 corresponding to correctly received and lost packets, respectively.

4.1 Renewal Error Process

For a renewal error process, the lengths of successive gaps are independent and distributed according to a common distribution [23]. Let $p(i)$ denote the probability that exactly $i-1$ 0s follow a 1, i.e.,

$$p(i) = Pr(0^{i-1}1|1), \tag{3}$$

where 0^{i-1} is a shorthand for $i-1$ successive 0's. Let $P(i)$ denote the probability that at least $i-1$ 0s follow a given error, i.e.,

$$P(i) = Pr(0^{i-1}|1). \tag{4}$$

Similarly, let $q(i)$ denote the probability that exactly $i-1$ 1s follow a 0, i.e., $q(i) = Pr(1^{i-1}0|0)$, and let $Q(i)$ denote the probability that at least $i-1$ 1s follow a 0, i.e.

$$Q(i) = Pr(1^{i-1}|0). \tag{5}$$

Let π_1 (π_0) denote steady-state probability of 1 (0). $R(m,n)$ denotes the probability that exactly $m-1$ 1s (errors) occur in the next $n-1$ bits following an error can be easily computed by the following recursion:

$$R(m,n) = \begin{cases} p(n), & \text{for } m = 1 \text{ and } n \geq 1, \\ \sum_{s=1}^{n-m+1} p(i)R(m-1, n-i), & \text{for } 2 \leq m \leq n. \end{cases}$$

4.2 Calculation of Residual Packet Loss Rate

A packet n cannot be recovered after decoding, if both packets n and $n+r$ are lost. Thus, to calculate the residual packet loss rate, it is sufficient to find the probability of the event where we have two 1s separated by $r-1$ bits each of which can be either a 1 or 0, as shown in Fig. 13. The probability of the first 1 is given by π_1. The $r-1$ bits that follow this 1 can have any number of 0s. Thus, we sum up

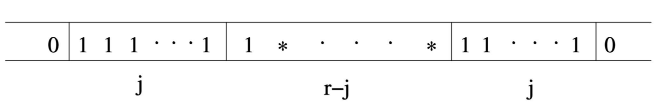

Fig. 14. An illustration of Case 1.

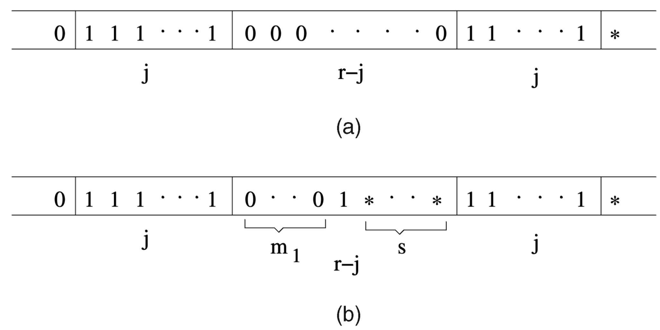

Fig. 15. An illustration of Case 2.

the probabilities, $R(m,r)$, that exactly $m-1$ errors occur between the 1s, for $1 \leq m \leq r$. So, the Residual Packet Loss Rate ($RPLR$) is given by

$$RPLR(p,q,r) = \sum_{m=1}^{r} \pi_1 \times R(m,r). \tag{6}$$

4.3 Calculation of Average Burst Length

When we say that a maximum burst of i packets have been lost, it means that both the original packets as well as their (FEC) copies are lost. Let p_j denote the probability that the burst length is j, which excludes that $j+1$ packets are lost in a sequence. We have two cases depending on the values of r and j.

4.3.1 The Bursts of the Original and the Copy Packets Do Not Overlap, i.e., $(j-1) < r$

In order to obtain an expression for p_j in this case, we consider the cases when the burst length is j. The burst length is j (but not $j+1$), in the following cases. The sequence of original and corresponding copy packets are both lost and one of the following happens:

Case 1. The packet after (before) the burst of original packets is lost (not lost) and the packet after the corresponding burst of copy packets is not lost, as depicted in Fig. 14. We denote the probability of this case as $p_{j,01*0}$.

Case 2. The packets before and after the burst of original packets are not lost, as depicted in Fig. 15. We denote the probability of this case as $p_{j,00**}$.

Case 3. The packets before and after the burst of original packets are lost and the packets before and after the corresponding burst of copy packets are not lost, as depicted in Fig. 16. We denote the probability of this case as $p_{j,1100}$.

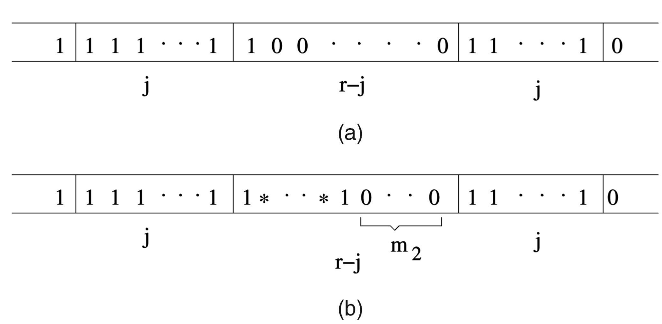

Fig. 16. An illustration of Case 3.

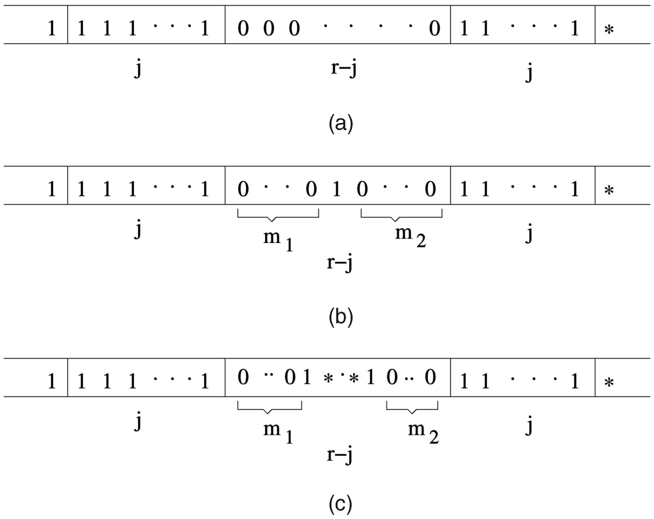

Fig. 17. An illustration of Case 4.

Case 4. The packet before (after) the burst of original packets is lost (not lost) and the packet before the corresponding burst of copy packets is not lost, as depicted in Fig. 17. We denote the probability of this case as $p_{j,100*}$. Summing up all the above probabilities, we have $p_j = p_{j,01*0} + p_{j,00**} + p_{j,1100} + p_{j,100*}$. We now proceed to give expressions for each of the above probabilities. In the first case (Case 1), as shown in Fig. 14, we have a 0 followed by $j+1$ 1s followed by $r-j-1$ bits, which can be either a 0 or a 1 followed by j 1s and ending with a 0. The probability of the first 0 is given by π_0. Then, the probability of the first one following it is given by $Q(2)$. The probability of no zeros in the next j bits is given by $R(j,j)$. The next $r-j-1$ bits followed by a 1 can have any number, $t-1$, of errors, the probability of which is given by $R(t, r-j)$, for $1 \le t \le r-j$. The next $j-1$ bits are 1s which are followed by a 0, the probability of which is given by $R(j-1, j-1)$ and $P(2)$, respectively. Thus, it follows that

$$p_{j,01*0} = \pi_0 \times Q(2) \times R(j,j) \times \left[\sum_{t=1}^{r-j} R(t,(r-j))\right] \times R(j-1,j-1) \times P(2).$$

In the second case (Case 2), as shown in Fig. 15, we have a 0 followed by j 1s followed by $r-j$ bits where the first bit has to be 0 and the rest can be either 0 or 1. This is then followed by j 1s. The probability of the first 0 is given by π_0. Then, the probability of the first one following it is given by $Q(2)$.

The probability of no zeros in the next $j-1$ bits is given by $R(j-1,j-1)$. The next $r-j$ bits following the last 1 has to be a 0 followed by $r-j-1$ arbitrary bits. In order to calculate the probability for these $r-j$ bits, we divide this case further into two subcases. In the first subcase, all these $r-j$ are 0s. The probability of which is given by $R(1, r-j+1)$. In the second subcase, these $r-j$ bits are composed of m_1 0s, $1 \le m_1 \le r-j-1$, followed by a 1 and $s = r-j-(m_1+1)$ arbitrary bits. The probability of the m_1 0s followed by a 1 is given by $R(1, m_1+1)$. The probability of next s bits can be calculated similarly to Case 1 shown above, by considering $t-1$ errors in them, for all $1 \le t \le r-j-m_1$. The next j bits are 1s, the probability of which is given by $R(j-1,j-1)$. Thus

$$\begin{aligned} p_{j,00**} &= \pi_0 \times Q(2) \times R(j-1,j-1) \\ &\times \left[\sum_{m_1=1}^{r-j-1}\sum_{t=1}^{r-j-m_1} R(1,t+1)R(t,r-j-m_1) \right. \\ &\left. + R(1,r-j+1)\right] \times R(j-1,j-1). \end{aligned}$$

In the third case (Case 3), as shown in Fig. 16, we have a 1 followed by $j+1$ 1s followed by $r-j-1$ bits where last bit has to be 0 and the rest can be either 0 or 1. This is then followed by j 1s and a 0.

The probability of the first 1 is given by π_1. The probability of no 0s in the next $j+1$ bits is given by $R(j+1,j+1)$. The next $r-j-1$ bits following the last 1 has to be $r-j-2$ arbitrary bits followed by a 0. In order to calculate the probability for these $r-j-1$ bits, we divide this case further into two subcases. In the first subcase, all these $r-j-1$ are 0s. The probability of which is given by $R(1, r-j)$. In the second subcase, these $r-j-1$ bits start with $s = r-j-1-(m_1+1)$ arbitrary bits and end with a 1 and m_2 0s, $1 \le m_1 \le r-j-2$. The probability of s bits can be calculated similarly as above, by considering $t-1$ errors in them, for all $1 \le t \le r-j-m_2-1$. The probability of following m_2 0s is given by $R(1, m+2+1)$. The next $j+1$ bits are 1s followed by a 0, the probability of which is given by $R(j-1,j-1) \times P(2)$. Thus

$$\begin{aligned} p_{j,1100} &= \pi_1 \times R(j+1,j+1) \\ &\times \left[R(1,r-j) + \sum_{m_2=1}^{r-j-2}\sum_{t=1}^{r-j-m_2-1} R(t,r-j-1-m_2) \right. \\ &\left. \times R(1,m_2+1)\right] \times R(j-1,j-1) \times P(2). \end{aligned}$$

In the last case (case 4), as shown in Fig. 17, we have $j+1$ 1s followed by $r-j$ bits where the first and last bits are 0. This is then followed by j 1s. The probability of the first 1 is given by π_1. The next j bits are 1s, the probability of which is given by $R(j,j)$.

To calculate the probability of the next $r-j$ bits, we divide this case into three subcases. In the first subcase, all these $r-j$ bits are 0s, the probability of which is given by $R(1, r-j+1)$. In the second subcase, these $r-j$ bits consist of m_1 0s, $1 \le m_1 \le r-j-2$, followed by a 1 and the rest of the bits being 0s, the probability being $R(1, m_1+1) \times R(1, r-j-m_1)$ for a particular m_1. In the final subcase, these $r-j$ bits consist of m_1 0s, $1 \le m_1 \le r-j-3$, followed by a 1, which is followed by a number of arbitrary bits and a 1 and m_2 0s, $1 \le m_2 \le r-j-m_1-2$. The probability of the first m_1 and the last m_2 0 bits along with the following 1s is given by $R(1, m_1+1)$ and $R(1, m_2+1)$, respectively. The probability of the $s = r-j-1-m_1-m_2$ arbitrary bits can be calculated as below by considering

0	1 1 1 1 · · · 1 1 1 1	0

r−1

Fig. 18. An illustration for $j-1 \geq r$.

$t-1$ errors, $1 \leq t \leq r-j-l-m_1-m_2$. Hence, finally we get

$$\begin{aligned} p_{j,100*} &= \pi_1 \times R(j,j) \\ &\quad \times \Bigg[R(1, r-j+1) + \sum_{m_1=1}^{r-j-2} R(1, m_1+1) \\ &\quad + \sum_{m_1=1}^{r-j-3} \sum_{m_2=1}^{r-j-m_1-2} R(1, m_1+1) \\ &\quad \times \left(\sum_{t=1}^{r-j-1-m_1-m_2} R(t, r-j-1-m_1-m_2) \right. \\ &\quad \left. \times R(1, m_2+1) \right) \Bigg] \times R(j-1, j-1). \end{aligned}$$

4.3.2 The Bursts of the Original and the Copy Packets Overlap, i.e., $(j-1) \geq r$

In this case, as shown in Fig. 18, we have a 0 followed by a sequence of $r+j$ 1s and a 0. The probability of the first 0 is given by π_0. The first 1 that occurs after this 0 occurs with a probability of $Q(2)$. The probability of the next $r+j-1$ 1s and the following 0 is given by $R(r+j-1, r+j-1)$ and $P(2)$, respectively. Thus,

$$p_j = \pi_0 \times Q(2) \times R(r+j-1, r+j-1) \times P(2).$$

From the above two cases, the average burst length is given by

$$ABL(p,q,r) = \sum_{j=1}^{n} \frac{j \times p'_j}{n}, \tag{7}$$

where

$$p'_j = \frac{p_j}{\sum_{j=1}^{n} p_j} \quad \text{and}$$

$$p_j = \begin{cases} p_{j,01*0} + p_{j,00**} + p_{j,1100} + p_{j,100*}, & \text{if } (j-1) < r, \\ \pi_0 \times Q(2) \times R(r+j-1, r+j-1) \\ \quad \times P(2), & \text{otherwise.} \end{cases}$$

4.4 Computing the Values of $p(i)$, $P(i)$, and $Q(i)$ for GE Loss Process

The GE error model, as explained in Section 3.3, is a two-state Markovian model and can be used to describe the temporal behavior of packet losses on a link [22], [24]. It has two states known as good state (or state 0) and bad state (or state 1). Good and bad states denote the correct reception and loss of a packet, respectively. p and q are the transition probabilities between the states. The residence times for states 0 and 1 are both geometrically distributed with means $\frac{1}{p}$ and $\frac{1}{q}$, respectively. The probability that n consecutive packets are lost is equal to $(1-q) \times q^{n-1}$, and thus, the residence time for state 1 is geometrically distributed.

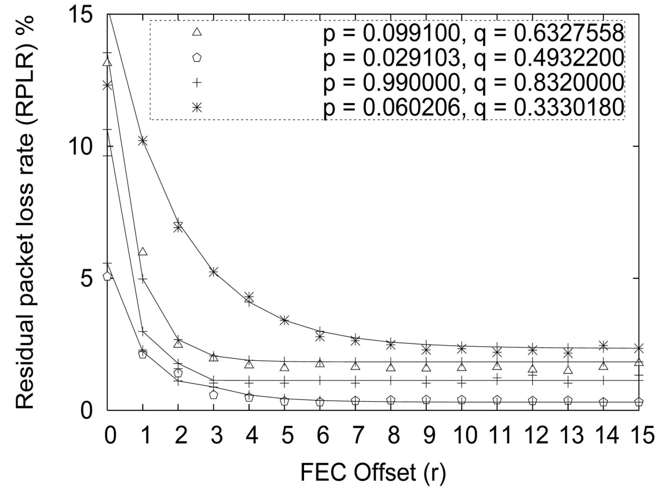

Fig. 19. Residual packet loss rate (percent) (lines correspond to analytical values and points correspond to simulation values).

Assuming that the channel loss process can be characterized by the GE model, we can compute the values of equations (given in Section 4.1) $p(i)$ (3), $P(i)$ (4), and $Q(i)$ (5). Recall that for the GE model the steady-state probabilities are given by

$$\pi_1 = \frac{p}{p+q}, \qquad \pi_0 = \frac{q}{p+q},$$

and computing the values of $p(i)$, $P(i)$, and $Q(i)$ in terms of p and q, we have

$$p(i) = \begin{cases} 1-q, & \text{if } i=1, \\ q*(1-p)^{i-2}*p, & \text{otherwise,} \end{cases}$$

$$P(i) = \begin{cases} 1, & \text{if } i=1, \\ q*(1-p)^{i-2}, & \text{otherwise,} \end{cases}$$

$$Q(i) = \begin{cases} 1, & \text{if } i=1, \\ p*(1-q)^{i-2}, & \text{otherwise.} \end{cases}$$

4.5 Validation

We evaluated the correctness of (6) and (7) by measuring $RPLR$ and ABL for different network loss rates through simulations in the NS-2 network simulator. As already mentioned, we have used the GE two-state Markov error model [22] available in the NS-2 for each link in our experiments. The simulation parameters are the same as shown in Table 1. By setting the good and bad state probabilities in the GE two-state Markov error model, we have simulated the bursty packet losses. We plot $RPLR$ and ABL against $FEC\text{-}Offset$ for different network losses (given in terms of network loss parameters p and q), as shown in Figs. 19 and 20. The analytical values were calculated from network loss parameters p and q, and these values (represented by lines in Figs. 19 and 20) closely match with simulation results (represented by points in Figs. 19 and 20) which validate the correctness of $RPLR$ and ABL equations (see Figs. 19 and 20). We used Mean Square Error (MSE) metric to quantitatively measure the error between analytical results and simulation results. The MSEs calculated in

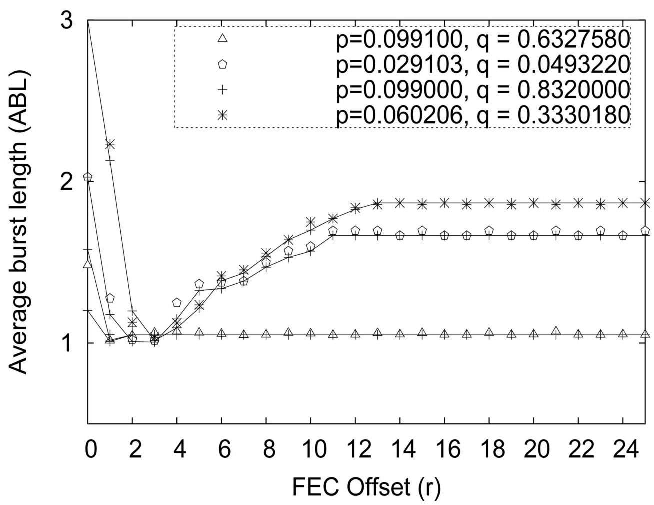

Fig. 20. Average burst length (lines correspond to analytical values and points correspond to simulation values).

Fig. 19 are as follows: 0.641 ($p = 0.099100$, $q = 0.6327558$), 0.590 ($p = 0.029103, q = 0.4932200$), 0.563 ($p = 0.990000, q = 0.8320000$), and 0.294 ($p = 0.060206$, $q = 0.3330180$). Additionally, the MSEs calculated in Fig. 20 are 0.244, 0.461, 0.322, and 0.381. From the graphs, we make two important observations. First, the $RPLR$ decreases with increasing $FEC\text{-}Offset$, r value, and second, the ABL decreases till it reaches an optimal $FEC\text{-}Offset$, r value after which it again increases. Hence, selecting $FEC\text{-}Offset$, r value by considering the point where ABL is minimum gives the best voice quality.

5 Performance Evaluation

We simulated our scheme using the NS-2 network simulator. The simulation parameters are shown in Table 1. All the simulation setup details are the same as given in Section 3.1. The background CBR sources generate 30 fixed size packets (1,500 bytes) per second. The number of background flows considered are 10. We modified the DSR [20] protocol to obtain two maximally node-disjoint routes from source to destination. For all the schemes, a voice packet is considered to be lost if redundant copies, both original and reconstructed (after FEC recovery), of that packet are lost. While varying mobility, the ABT (the average time spent in the bad state) is kept constant at 30 ms. For traffic other than voice traffic, the default MAC layer behavior is considered, and acknowledgments are not turned off. Acknowledgments are turned off only for voice traffic. We explain the voice packet flow in the network as follows: On receiving a voice frame from the higher layer, the IP makes use of precedence (priority) field in the IP header to mark it as the voice packet. This field helps IP to distinguish voice packets from other kinds of traffic. However, this information is not available to the MAC layer. To provide such information to the MAC layer, we need support from the network layer protocol (i.e., IP). When the network layer passes a voice packet to the MAC queue, it extracts the value of precedence field from the IP header and passes to the MAC layer. This will allow the MAC layer to distinguish a voice packet from other type of packets (control, background, etc.).

5.1 Validation of the Packetization Scheme

Simulation results (FLR and average burst size) are measured at the application layer of the destination node by counting the actual number of voice frames received after decoding. For validating these simulation results, we measured the actual network loss parameters p and q for a given flow at the network layer of destination node (i.e., before FEC recovery). By substituting these network loss parameters p, q in (6) and (7), we theoretically determined the $RPLR$ and ABL. Figs. 21a and 21b show the changes in FLR for varying mobility and ABT, respectively. Figs. 21c and 21d show the changes in average burst size for varying mobility and ABT, respectively. As observed from the graphs, simulation results closely match with analytical results. The MSEs computed in Figs. 21a, 21b, 21c, and 21d are 0.23, 0.54, 0.26, and 0.29, respectively. Our analytical model is generic such that it depends only on packet losses (induced by mobility, channel fading, collisions, path breaks, etc.).

5.2 Packetization Scheme versus Layered Scheme versus MD Scheme

We compare the following two schemes with our scheme to evaluate the effectiveness of our proposed packetization scheme. The 802.11 DCF with 0-MAC retransmissions is used in all the schemes to ensure fairness while comparing them.

- **Layered Scheme**. In this scheme, the raw voice stream is encoded into the BL substream and one EL substream using a layered codec [2]. Multipath transport is used to transmit the BL substream along one path and the EL substream along another path. Both the BL and EL packets carry a FEC copy of previous BL packet. The voice payload sizes (shown in Table 3) for layered scheme are calculated after adding FEC copy (BL packet), sequence number, and timestamp bits to the BL and EL bit streams shown in Fig. 7b.
- **MD Scheme**. In this scheme, using MD coding, the raw voice stream is encoded into two descriptions [10]. Both the descriptions are protected using FEC with offset 1. The voice payload size of each description is 42 bytes after adding sequence number and timestamp bits (shown in Table 3) [3].

The voice payload sizes of the packetization scheme are computed in Fig. 7a and are shown in Table 3. For all simulations, the $delay_{threshold} = 200$ ms. The voice packet size in packetization scheme is adaptive, and it varies from 28 to 36 bytes. If at least the BL (in case of layered scheme and in packetization scheme) or one of the two descriptions (in case of MD scheme) is received, then that voice frame is assumed to be received by the destination successfully. The overall packet overhead is low in case of packetization scheme as shown in the last column of Table 3. Thus, without increasing the packet overhead, our scheme gives the best performance over other schemes.

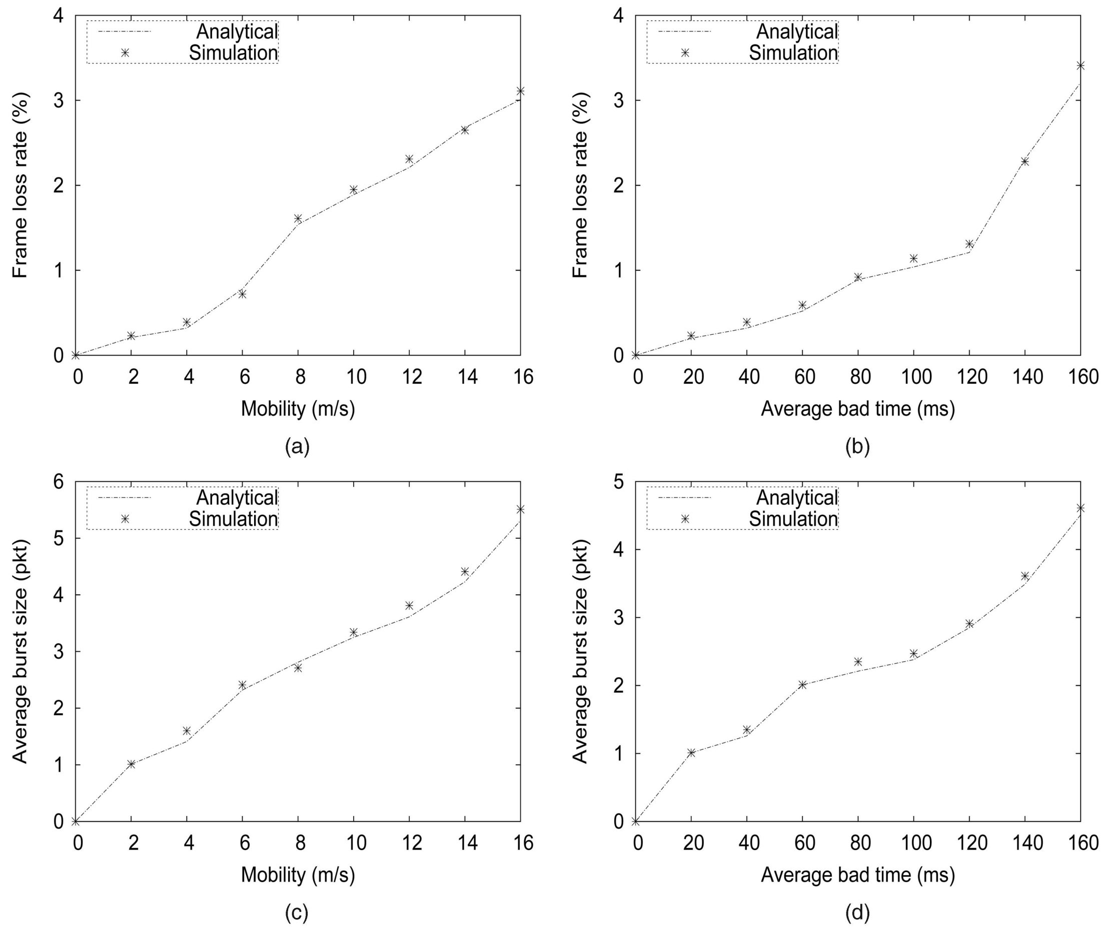

Fig. 21. Comparison between simulation and analytical results of the packetization scheme. (a) FLR versus mobility. (b) FLR versus ABT. (c) Average burst size versus mobility. (d) Average burst size versus ABT.

TABLE 3
Packet Overheads in Various Schemes

Scheme	Path 1 Voice Payload Size (bytes)	Path 2 Voice Payload Size (bytes)	Total Payload Size (bytes)	Overhead due to Lower Layers (UDP+IP+802.11 MAC) over Two Paths (bytes)	Total Packet Size over Two Paths (bytes)
Layered Scheme	34	48	34+48=82	2 * (8+20+24)=104	186
MD Scheme [3]	46	46	46+46=92	2 * (8+20+24)=104	196
Packetization Scheme	28/36	28/36	36+36=72	2 * (8+20+24)=104	176

5.2.1 Effect of ABT

Fig. 22 shows the changes in *FLR* for various schemes by varying ABT. The *FLR* of layered scheme increases rapidly, when the ABT increases beyond 40 ms. Within the allowed delay, if the BL packet of layered scheme is not received, then the layered scheme tries to recover it through its FEC copies present in the succeeding packets. On the other hand, if the scheme receives just EL packet, it is not useful for decoding and it still has to wait for succeeding packet to get its BL bits causing more delays and, thus, more packet losses. MD scheme performs better than layered scheme. However, under high ABT, the packetization scheme outperforms the MD scheme. As the packetization scheme adaptively adjusts its $FEC\text{-}Offset$ based on the network state, it is able to sustain burst losses better than the MD scheme. The MD scheme always uses $FEC\text{-}Offset = 1$, and thus, it cannot sustain bursty losses compared to the packetization scheme.

Fig. 23 shows the changes in *end-to-end delay* under ABT for various schemes. The *end-to-end delay* in layered scheme is higher than in the other schemes as it increases the overall network traffic causing more collisions and more packet losses. Our packetization scheme performs better than MD scheme as it uses adaptive FEC and variable packet size, thus resulting in overall reduction in both packet overhead and network traffic.

Fig. 24 shows the average size of bursts for various schemes. When the channel is less bursty (i.e., average bad state dwell time is under 20 ms), random loss is the dominant source of loss, and only MD scheme and packetization scheme can sustain packet losses. When the channel gets bursty, the average burst size of layered scheme is relatively high because most of the BL packets are lost either due to channel burstiness (causing more delays to receive BL FEC copy). MD scheme has better performance compared to layered scheme, but the average burst

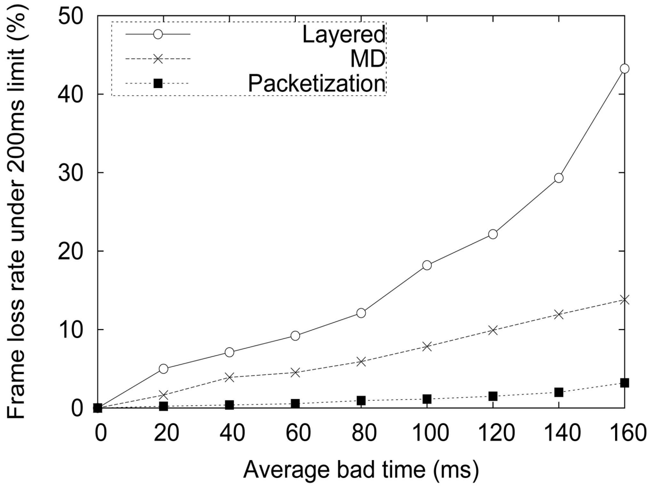

Fig. 22. FLR versus ABT under 200-ms threshold.

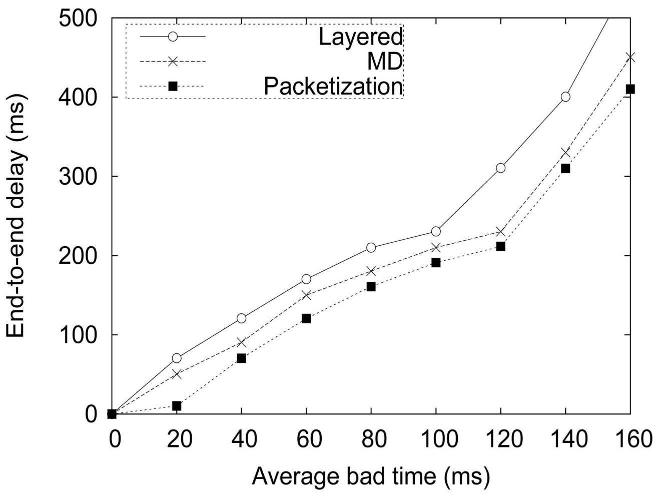

Fig. 23. End-to-end delay versus ABT.

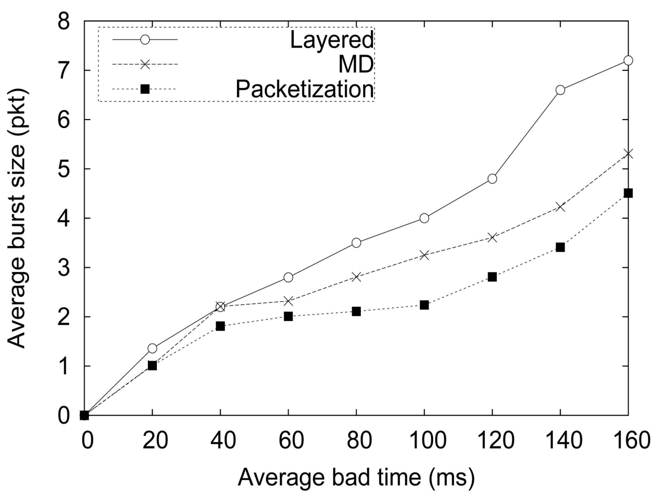

Fig. 24. Average burst size versus ABT.

size in the MD scheme is still higher especially when average bad state dwell time is greater than 40 ms. However, our packetization scheme employs adaptive FEC, and hence, it performs better than both layered and MD schemes.

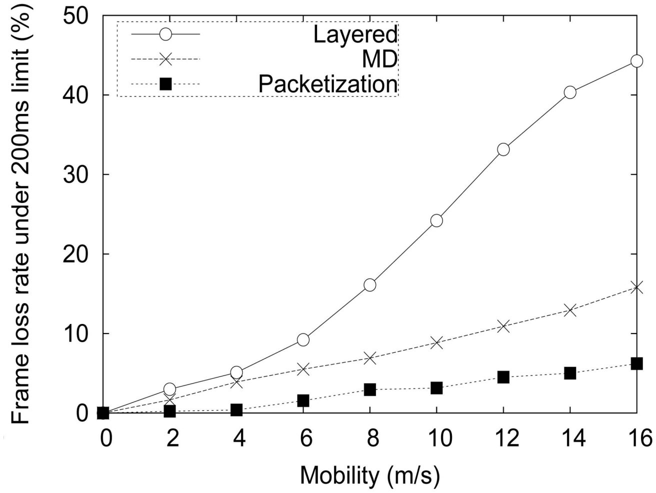

Fig. 25. FLR versus mobility under 200-ms threshold.

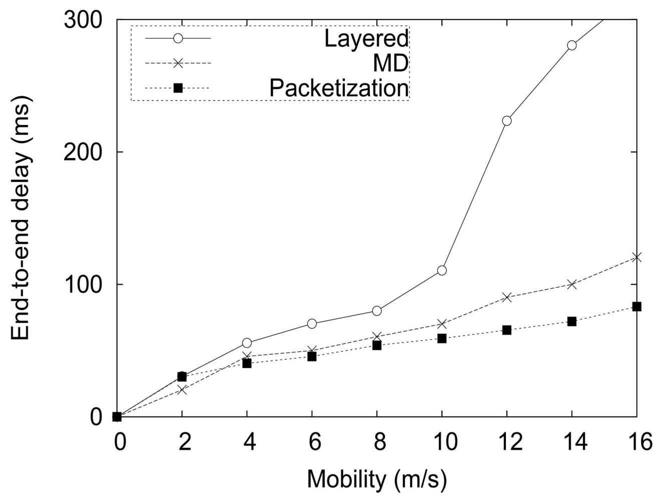

Fig. 26. End-to-end delay versus mobility.

5.2.2 *Effect of Mobility*

Fig. 25 shows the overall *FLR* for varying mobility with a threshold delay of 200 ms. As shown in the figure, the loss rates of layered scheme and MD scheme are higher compared to packetization scheme though MD scheme performs better than layered scheme. It is because the packetization scheme employs FEC adaptively, thereby having better chances of sustaining bursty packet losses compared to MD scheme. Fig. 26 shows the *end-to-end delay* for varying mobility for all the three schemes. As observed from the figure, our packetization scheme performs better than the other two schemes.

5.3 Measurement of Perceptual Evaluation Speech Quality Mean Opinion Score (PESQ-MOS)

At the destination, voice frames are decoded and the Wide Band (WB) version of ITU perceptual measurement algorithm, PESQ-MOS reference software tool [25], is used to measure the *perceived voice quality*. The PESQ-MOS tool compares the degraded speech with the reference speech and gives the objective MOS value in a five-point score ranging from −0.5 (worst) to 4.5 (best). Using the PESQ-MOS tool, we evaluated the voice quality score for various voice packets. For this evaluation, we used raw voice frame

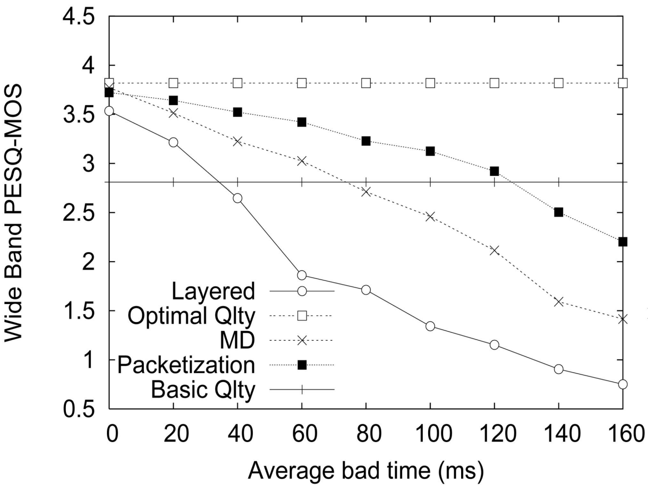

Fig. 27. WB PESQ-MOS versus ABT.

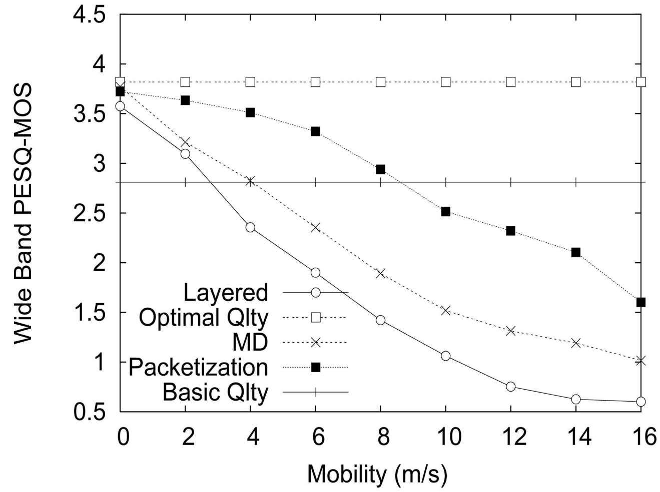

Fig. 28. WB PESQ-MOS versus mobility.

as the reference input. If we give raw voice frame also as the test input, the PESQ-MOS tool gives a score of 4.5. Likewise, if we give the decoded bits of AMR-WB voice frame as the test input to the PESQ-MOS tool, it gives a score of 3.818. It is to be noted that there is a decrease of PESQ-MOS score from 4.5 to 3.818 for the decoded AMR-WB voice frame. This is due to a lossy compression method used by the AMR-WB coder. Hence, the optimal PESQ-MOS that can be obtained at the destination node, assuming that there are no losses in the network, is 3.818. If we give the BL decoded bits of AMR-WB voice frame of packetization scheme as test input, it gives a score of 2.81 which corresponds to the basic quality. If we give the decoded bits BL and description 1 or description 2 (MDC1/MDC2) of EL of AMR-WB voice frame of packetization scheme as test input to PESQ-MOS tool, it gives a score of 3.24. This means that, in the packetization scheme, if only one of the two voice packet entities is successfully received, then the destination node obtains a voice quality of 3.24. If one of the two descriptions of AMR-WB voice frame in the case of MD scheme is given as the test input, it gives a score of 2.86. Fig. 27 shows the measured WB PESQ-MOS at the destination for varying ABT. Since the BL is protected strongly in the packetization scheme and at all the times at least BL substream, BL_i, and $MDC1_{EL(i)}$ and/or $MDC2_{EL(i)}$ of EL substream are available at the destination node, it performs better than the other schemes. This quality is better than the quality obtained in the case of MD scheme, in which only one of the two descriptions (which is equal to BL quality) is received in case of packet loss in the network. Fig. 28 shows the measured WB PESQ-MOS at the destination for varying mobility. The packetization scheme exhibits better quality compared to both layered and MD schemes. The performance improvement of the packetization scheme is due to the fact that it makes use of variable packet size (due to adaptive FEC mechanism) and it better utilizes the bandwidth for improving the *perceived voice quality*. Both optimal and basic quality scores are also shown in Figs. 27 and 28. Thus, our scheme provides the best quality with minimal overhead at all the times compared to the existing schemes.

6 Conclusions

By exploiting the strengths of adaptive FEC, layered coding, and MD coding, we proposed an effective packetization scheme to achieve the best *perceived voice quality* while not increasing the overheads associated in transmitting small-sized voice packets and, thus, making it feasible to deploy voice application over AWNs. We measured the perceived speech quality for different schemes by integrating the NS-2 network simulator with a real adaptive speech codec (AMR-WB codec) and a perceived quality evaluation system based on the WB version of ITU-T PESQ. We presented an analytical approach based on renewal error process to calculate the *residual packet loss rate* and *average burst length* after error recovery. We then computed these parameters for a GE loss process and compared with experimental data. In our study, the two-state GE model is used in both the algorithm's model of the channel as well as in the channel itself to create errors in the simulation. As part of our future work, we plan to carry out the following: 1) evaluate the performance of our packetization scheme by considering other channel error models in the simulation, 2) study the effect of nonreal time traffic on the real-time traffic, and 3) demonstrate the effectiveness of our scheme using an experimental testbed.

Acknowledgments

The authors wish to thank G.V. Shyamnath and the anonymous reviewers for their valuable comments and suggestions. This work was supported by Microsoft Research University Relations, India.

References

[1] ITU-T Recommendation, G.114, *One Way Transmission Time,* Feb. 1996.
[2] J.D. Gibson, A. Servetti, H. Dong, A. Gersho, T. Lookabaugh, and J.C. De Martin, "Selective Encryption and Scalable Speech Coding for Voice Communications over Multihop Wireless Links," *Proc. IEEE Military Comm. Conf. (MILCOM '04),* vol. 2, pp. 792-798, Nov. 2004.

[3] C.-h. Lin, H. Dong, U. Madhow, and A. Gersho, "Supporting Real-Time Speech on Wireless Ad Hoc Networks: Inter-packet Redundancy, Path Diversity, and Multiple Description Coding," *Proc. Second ACM Int'l Workshop Wireless Mobile Applications and Services on WLAN Hotspots (WMASH '04),* pp. 11-20, Oct. 2004.
[4] X. Yu, J.W. Modestino, and I.V. Bajic, "Modeling and Analysis of Multipath Video Transport over Lossy Networks," *Proc. 11th Int'l Conf. Distributed Multimedia Systems (DMS '05),* pp. 265-270, Sept. 2005.
[5] L. Munoz, M. Garcia, J. Choque, R. Aguero, and P. Mahonen, "Optimizing Internet Flows over IEEE 802.11b Wireless Local Area Networks: A Performance-Enhancing Proxy Based on Forward Error Correction," *IEEE Comm. Magazine,* vol. 39, no. 12, pp. 60-67, Dec. 2001.
[6] S. Aramvith, C.-W. Lin, So. Roy, and M.-T. Sun, "Wireless Video Transport Using Conditional Retransmission and Low-Delay Interleaving," *IEEE Trans. Circuits and Systems for Video Technology (CSVT '02),* vol. 12, no. 6, pp. 558-565, June 2002.
[7] G. Rubino and M. Varela, "Evaluating the Utility of Media-Dependent FEC in VoIP Flows," *Proc. Fifth Int'l Workshop Quality of Future Internet Services (QofIS '04),* pp. 31-43, Sept. 2004.
[8] Z. Zhou, P.K. McKinley, and S.M. Sadjadi, "On Quality-of-Service and Energy Consumption Tradeoffs in FEC-Encoded Audio Streaming," *Proc. 12th IEEE Int'l Workshop Quality of Service (IWQoS '04),* pp. 161-170, June 2004.
[9] H. Wu, C. Hung, M. Gerla, and R. Bagrodia, "Speech Support in Wireless Multihop Networks," *Proc. IEEE Int'l Symp. Parallel Architectures, Algorithms, and Networks (ISPAN '97),* pp. 282-288, Dec. 1997.
[10] H. Dong, A. Gersho, J. Gibson, and V. Cuperman, "A Multiple Description Speech Coder Based on AMR-WB for Mobile Ad Hoc Networks," *Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing (ICASSP '04),* vol. 1, pp. 277-280, May 2004.
[11] Y. Wang, M. Orchard, V. Vaishampayan, and A. Reibman, "Multiple Description Coding Using Pairwise Correlating Transforms," *IEEE Trans. Image Processing,* vol. 10, no. 3, pp. 351-366, Mar. 2001.
[12] V.K. Goyal and J. Kovacevic, "Generalized Multiple Description Coding with Correlating Transforms," *IEEE Trans. Information Theory,* vol. 47, no. 6, pp. 2199-2224, Sept. 2001.
[13] J.G. Apostolopoulos, T. Wong, W. Tan, and S. Wee, "On Multiple Description Streaming with Content Delivery Networks," *Proc. IEEE INFOCOM '02,* vol. 3, pp. 1736-1745, June 2002.
[14] J-C. Bolot, S. Fosse-Parisis, and D. Towsley, "Adaptive FEC-Based Error Control for Internet Telephony," *Proc. IEEE INFOCOM '99,* vol. 3, pp. 1453-1460, Mar. 1999.
[15] C. Perkins, O. Hodson, and V. Hardman, "A Survey of Packet Loss Recovery Techniques for Streaming Audio," *IEEE Network,* vol. 12, no. 5, pp. 40-48, Oct. 1998.
[16] Y.S. Liaw, A. Dadej, and A. Jayasuriya, "Throughput Performance of Multiple Independent Paths in Wireless Multihop Network," *Proc. IEEE Int'l Conf. Comm. (ICC '04),* vol. 7, pp. 4157-4161, June 2004.
[17] *The Network Simulator—NS-2,* http://www.isi.edu/nsnam/ns, 2008.
[18] T. Camp, J. Boleng, and V. Davies, "A Survey of Mobility Models for Ad Hoc Network Research," *Wireless Comm. and Mobile Computing: Special Issue on Mobile Ad Hoc Networking: Research, Trends and Applications,* vol. 2, no. 5, pp. 483-502, Sept. 2002.
[19] http://ghost.lesiuk.org/AdHoc/network_simulator/, 2008.
[20] D.B. Johnson and D.A. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks," *Mobile Computing,* vol. 353, pp. 153-181, Kluwer Academic Publishers, 1996.
[21] *3GPP TS 26.171: Adaptive Multi-Rate Wide Band Speech Codec,* http://www.3gpp.org/ftp/Specs/html-info/26171.htm, 2008.
[22] W. Jiang and H. Schulzrinne, "QoS Measurement of Internet Real-Time Multimedia Services," Technical Report-CUCS-015-99, Columbia Univ., Dec. 1999.
[23] E. Elliot, "A Model of the Switched Telephone Network for Data Communications," *Bell Labs Technical J.,* pp. 80-109, Jan. 1963.
[24] M. Yajnik, S. Moon, J. Kurose, and D. Towsley, "Measurement and Modelling of the Temporal Dependence in Packet Loss," *Proc. IEEE INFOCOM '99,* vol. 1, pp. 345-352, Mar. 1999.
[25] Proposed Modification to P.862 to Allow PESQ to be Used for Quality Assessment of Wideband Speech, *ITU-T SG12 Delayed Contribution COM-D007-E,* Feb. 2001.

Venkat Raju Gandikota received the BTech degree in computer science and engineering from Nagarjuna University, Guntur, India, in 2002 and the MS degree in computer science and engineering from the Indian Institute of Technology (IIT) Madras, Chennai, India, in 2007. He is currently a senior systems engineer at the Yahoo! Software Development India Pvt. Ltd., Bangalore, India. His research interests include multimedia transport in wireless networks.

Bheemarjuna Reddy Tamma received the BTech degree in computer science and engineering from Andhra University, Visakhapatnam, India, in 2000, the ME degree in computer science and engineering from the National Institute of Technology (NIT), Rourkela, India, in 2002, and the PhD degree in computer science and engineering from the Indian Institute of Technology (IIT) Madras, Chennai, India, in 2007. In his doctoral thesis, he focused on QoS provisioning and multimedia transport in ad hoc wireless networks. From January 2007 to March 2007, he was a senior project officer at IIT Madras. He is currently a postdoctoral researcher at CalIT2, Jacobs School of Engineering, University of California, San Diego. His research interests include QoS provisioning, multimedia transport, and traffic characterization in wireless networks. He is a member of the IEEE.

C. Siva Ram Murthy received the PhD degree from the Indian Institute of Science, Bangalore, India. He has been with the Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai, India, since 1988, where he is currently a professor. His research interests include parallel and distributed computing, real-time systems, lightwave networks, and wireless networks. He has published more than 130 papers in refereed international journals and more than 125 papers in refereed international conferences in these areas. He has coauthored the textbooks *Resource Management in Real-time Systems and Networks* (MIT Press, Cambridge, USA), *WDM Optical Networks: Concepts, Design, and Algorithms* (Prentice Hall, New Jersey, USA), and *Ad Hoc Wireless Networks: Architectures and Protocols* (Prentice Hall, New Jersey, USA). He is a fellow of the Indian National Academy of Engineering, an associate editor of the *IEEE Transactions on Computers*, and a subject area editor of the *Journal of Parallel and Distributed Computing*. He is a senior member of the IEEE.