On Supporting Real-time Speech over Ad hoc Wireless Networks†

G. Venkat Raju, T. Bheemarjuna Reddy, Shyamnath Gollakota, and C. Siva Ram Murthy
Department of Computer Science and Engineering
Indian Institute of Technology Madras, India, 600036
{gvraju, arjun, shyam}@cs.iitm.ernet.in, murthy@iitm.ac.in

***Abstract*— Providing real-time speech support over multi-hop ad hoc wireless networks is a challenging task. In order to make a voice application to be feasible over ad hoc wireless networks, the *perceived voice quality* must be improved while reducing the packet overhead. In this work, we propose various mechanisms to achieve this objective. Using these mechanisms, we propose an efficient packetization scheme, in which, the important sub-stream of the voice stream is protected adaptively with forward error correction (FEC) depending upon the network state and is transmitted over two disjoint paths. The less-important sub-stream of the voice stream is encoded into two descriptions, which are then transmitted over two maximally node disjoint paths. We derived an expression for estimating *residual packet loss rate, RPLR* for the given *FEC-Offset, r* (the distance between original voice frame and piggybacked redundant voice frame) and packet loss in the network. Our scheme adapts the *FEC-Offset* value (it chooses the *FEC-Offset* that minimizes *RPLR* as much as possible) based on loss rate feedback obtained from the destination. As observed from simulations, the proposed scheme achieves significant gains in terms of reduced *frame loss rate*, reduced *control overhead*, and minimum *end-to-end delay* and almost double the *perceived voice quality* compared to the existing approaches.**

I. 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 pre-existing infrastructure. In this paper, we concentrate mainly on providing voice support over AWNs, because it is a key application in many scenarios. The standard retransmission based strategies proposed in the literature are poorly matched to voice applications, because of timeliness and large overheads involved (especially in wireless networks) in transmitting small sized voice packets [1]. The unique characteristics of voice application such as, small payload size (typically 20 bytes) and timely arrival (typically for interactive voice communication, the *end-to-end delay* must be less than 200 *ms* [2]) of the packets at the destination, make it very challenging to deploy over AWNs.

A potentially promising approach to reduce the voice packet loss rate is to establish multiple paths between the source and destination of the session and to use speech coding schemes that take advantage of the existence of multiple paths. One such coding scheme is Multiple Description (MD) coding [3], in which a voice stream is encoded into multiple sub-streams (descriptions). The authors of [4] use MD coding with path diversity for supporting video applications in ad hoc wireless networks. In [1], a combination of inter-packet redundancy, MD coding, and path diversity was used to provide speech support over ad hoc networks. However, there is no prior work that integrates design concepts across multiple layers to provide an integrated system suitable for real-time voice applications over ad hoc wireless networks. In this work, we suggest adaptive FEC in tandem with multi-path transport to reduce the *frame loss rate*. Use of adaptive FEC also helps in reducing the packet overhead. At the MAC layer, we avoid retransmissions (hence no acknowledgments (ACKs)) to minimize *control overhead* which also helps in reducing the *end-to-end delay*. We exploit the combined strengths of layered coding and MD coding in order to improve the *perceived voice quality*. Using the above proposed techniques, we propose a scheme which effectively combines adaptive FEC, layered coding, and MD coding to achieve significant performance gains in terms of *perceived voice quality*, *frame loss rate*, and *end-to-end delay*. As we explain later (in Section II), best *perceived voice quality* can be achieved as the scheme is more immune to path breaks and congestion in the network. The rest of the paper is organized as follows. Section II provides a detailed description of the proposed effective packetization scheme. Sections III and IV discuss the analytical framework and simulation results, respectively. Finally, section V contains concluding remarks.

II. AN EFFICIENT PACKETIZATION SCHEME

During voice communication, if the number of lost voice packets is higher than that tolerated by the listener, either an error control or loss recovery mechanism is required. Automatic Repeat Request (ARQ) mechanisms are closed-loop mechanisms where the source retransmits lost packets as reported by the destination. ARQ mechanisms 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. Media independent FEC schemes are not well suited to 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 non negligible *block delay* to the *end-to-end delay*, thereby decreasing the quality of interactivity between the participants [5]. We thus use media specific FEC for providing protection to the voice stream in our scheme. Media specific schemes piggyback information about the voice packets that correspond to present period with later voice packets as shown in Fig. 1. By increasing the amount of piggyback information added to

†This work was supported by the Department of Science and Technology, New Delhi, India.

0-7803-9746-0/06/$20.00©2006 IEEE

the voice stream, the media specific schemes recover from multiple losses. However, increasing the amount of piggyback information when the network loss rate is low will waste bandwidth. This motivates the need to develop methods to control the amount of redundancy depending on the network loss rate. A simple example use of this technique is to append the contents of the previous packet *n-1* into current packet *n* as shown in Fig. 1. Thus, even if packet *n-1* is lost, the data of packet *n-1* can be retrieved on receiving packet *n* at the destination. The *FEC-Offset* in this case is 1.

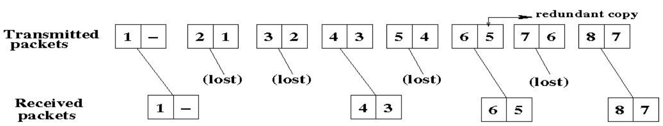

Fig. 1. **Typical FEC mechanism.**

Layered coding (LC) and MD coding have been two efficient scalable source coding approaches for multipath transmissions [1,3] and are robust against inevitable transmission errors. In contrast to a conventional media coder that generates a single bitstream, LC and MD coding based coders encode a media source into two or more sub-streams. For LC, one base layer (BL) sub-stream and one or more enhancement-layer sub-streams are generated. The BL sub-stream can be decoded to provide a basic quality (below which unacceptable) of voice while the enhancement layer (EL) sub-streams are mainly used to refine the quality of the voice, that is reconstructed from the base-layer sub-stream. If the base-layer sub-stream is lost, the EL sub-streams become useless, even if they are received perfectly. Thus, reliable transmission of base-layer sub-stream is necessary for LC to perform well.

However, for MD coding, these sub-streams, also called as descriptions can be decoded independently to produce a signal of basic quality and MD coding does not require prioritized transmission as all descriptions have equal importance. Since the probability of losing all descriptions is relatively low, it performs better than LC at higher packet loss rates. By exploiting the strengths of both LC and MD coding, we propose an effective packetization scheme at the application level as follows:

(i). Base layer sub-stream is protected using adaptive FEC and transmitted over two maximally node disjoint paths. Depending on the packet loss rate in the network, the *FEC-Offset* (see Fig. 1) is varied dynamically and thus, the BL is strongly protected. Due to this at least basic voice quality is ensured with minimum overhead always.

(ii). Enhancement layer sub-stream is encoded into two descriptions to take advantage of multipath transmission and to improve the perceived voice quality. Each path carries one of the two descriptions of the original EL sub-stream. When both the descriptions are received at the destination, the destination voice application can retrieve the original EL packet from the two descriptions. However, if only one of them is received, then the received description contributes in improving the quality of the BL sub-stream. Each transmitted packet contains the FEC protected BL and one of the two descriptions of the EL. Thus, with high probability both the BL sub-stream and one of the two descriptions of the EL sub-stream 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,

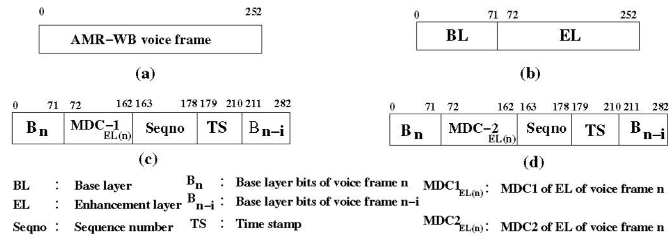

Fig. 2. **Packet formats (a) AMR-WB voice packet, (b) Layer wise bit classification, (c) Packet format for path 1, and (d) Packet format for path 2.**

we consider the Adaptive Multi-Rate Wide Band (AMR-WB) speech codec [6] with 12.65 Kbps bit rate. The multi-rate codec has eight encoding modes ranging from 6.6 Kbps to 23.85 Kbps. 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. 2(a). 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. 2(b). There are no Class C bits in this mode. The packet formats for paths 1 and 2 are shown in Figures 2(c) and 2(d), 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 Figures 2(c) and 2(d), respectively.

A. Working Mechanism of Packetization Scheme

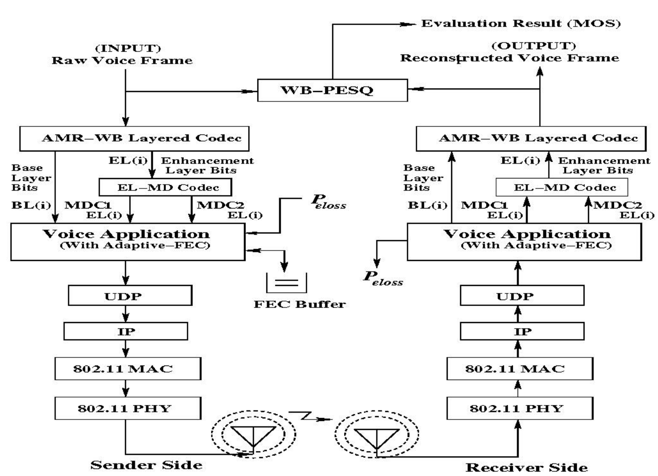

Fig. 3. **Layer wise description.**

As shown in Fig. 3, AMR-WB Layered Coder takes a raw voice frame as the input and produces the BL sub-stream and one EL sub-stream. The EL sub-stream 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 sub-stream and the corresponding two descriptions (Fig. 2(c) and 2(d)). 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. 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 of the two voice packet entities (correspond to a single voice frame) that come through two maximally node-disjoint paths are received correctly by the destination voice application, the EL-MD Codec, shown in the Fig. 3, re-combines the two EL descriptions of these two voice packet entities to get the original EL packet. Full voice quality is obtained in this case. If it receives only one voice packet entity, then the bits corresponding to the other description are made empty, before the EL packet is given to AMR-WB Layered Codec. A voice quality that is better than basic voice quality and less than full voice quality is obtained in this case. On the other hand if both voice packet entities are lost, then the voice application will use adaptive FEC mechanism to recover from the loss of voice frame. The AMR-WB Layered Codec combines the BL and the EL packets to obtain the original (or degraded) voice frame. The decoded voice frame will then be played by the voice application. In order the (FEC) redundant information to be most effective, the *FEC-Offset* should be varied dynamically based on the actual loss process in the network. The voice application at the source node should estimate the target perceived loss rate at the destination and it should choose the best *FEC-Offset* for sending the redundant voice data that will yield to the minimum loss rate. The source holds two parameters (*FEC Offset, FEC Buffer*) to serve the purpose of adaptive FEC mechanism. The *FEC Buffer* holds the recently transmitted voice frames temporarily so that FEC data can be added to the ongoing voice packets. The *FEC Offset* is used for adding FEC data adaptively along with the original voice frame. The source voice application module finds the best possible *FEC Offset*, r^* (as shown in Eqn. (1) below), after estimating *residual packet loss rate, RPLR* based on P_{eloss}, loss rate of the network. The next section deals with how we estimate *RPLR* given *FEC-Offset, r* and P_{eloss}, *i.e., RPLR(r,* P_{eloss}*)*.

$$r^* = \{r \ni RPLR(r, P_{eloss}) = min_{i=1}^{MAX\_FECOFFSET}\{RPLR(i, P_{eloss})\} \quad (1)$$

The term *MAX_FECOFFSET* refers to the maximum permissible *FEC-Offset* value in the network and $0 \leq r \leq$ *MAX_FECOFFSET*. By counting the number of received voice packets from the packet history, the destination node calculates the *packet loss rate*, P_{eloss}, present in the network. The destination node now feeds back this information to the source node through piggy-backed voice packets once the source and destination exchange their roles. That is, the sender node now becomes inactive and the destination node becomes active. The corresponding node then updates the *FEC Offset* (as shown in Eqn. (1)) to reflect the packet loss in the network. Since voice application is an interactive application, the overhead involved in exchanging packet loss parameters is almost negligible as the parameters are piggy-backed to the ongoing voice packets at both the source and destination side. Thus, the proposed scheme exploits both adaptive FEC (to protect BL sub-stream), MD coding (to transmit EL sub-stream), and multipath transport (to reduce correlated losses and to improve robustness) to provide better voice quality for voice communication over ad hoc wireless networks. For voice packets, if a packet is lost, then retransmitting that packet will not help in reducing the *end-to-end delay*. Further, retransmissions increase the network traffic. 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). If there are any lost packets in the voice stream, then the receiver will try to conceal the packet loss using the FEC mechanism.

III. Theoretical Estimation of Residual Packet Loss Rate

In this section, we derive analytical expressions for estimating the *residual packet loss rate* after the source node receives the end-to-end effective loss rate, P_{eloss} from the destination node. For convenience sake let P_{eloss} = p. We first calculate *RPLR* when *FEC-Offset* = 1 and then we extend it for *FEC-Offset* = *r*.

A. Calculation of *RPLR* when *FEC-Offset = 1*

Fig. 4 shows the case of protecting packets with *FEC-Offset = 1*. Consider blocks of losses such that each block is a continuous burst of packet losses. Suppose that we have s such blocks as shown in Fig. 5. Let the sizes of the blocks be $\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_s$ respectively, where

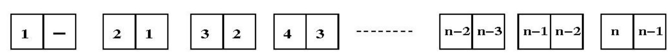

Fig. 4. **Packet transmission with *FEC-Offset* = 1.**

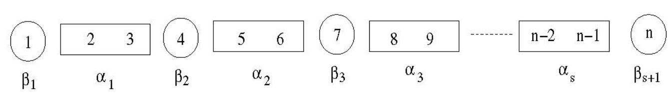

Fig. 5. **Dividing packets into two parts : burst packets and good packets.**

$$\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_s \quad \geq 1 \quad (2)$$

Let the blocks between them be $\beta_1, \beta_2, \beta_3, \cdots, \beta_s, \beta_{s+1}$ respectively, where

$$\begin{aligned} \beta_1 &\geq 0 \\ \beta_2, \cdots, \beta_s &\geq 1 \\ \beta_{s+1} &\geq 0 \end{aligned}$$

where α's and β's are integers. We have

$$\beta_1 + \beta_2 + \beta_3 + \cdots + \beta_{s+1} + \alpha_1 + \alpha_2 + \cdots + \alpha_s = n \quad (3)$$

From the values of β_i's, we have the following 4 cases: $\{\beta_{s+1} \geq 1$ and $\beta_1 \geq 1\}$, $\{\beta_{s+1} = 0$ and $\beta_1 \geq 1\}$, $\{\beta_{s+1} \geq 1$ and $\beta_1 = 0\}$, and $\{\beta_{s+1} = 0$ and $\beta_1 = 0\}$, we further take up each case in detail below.

Case(i): $\{\beta_{s+1} \geq 1$ **and** $\beta_1 \geq 1\}$: For every block of packets of size α_i, we lose $\alpha_i - 1$ packets after decoding if at least one packet follows this block α_i which is not lost. Therefore

$$\begin{aligned} (\alpha_1 - 1) + (\alpha_2 - 1) + (\alpha_3 - 1) + \cdots + (\alpha_s - 1) &= i \quad (or) \\ \alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_s &= i + s \quad (4) \end{aligned}$$

Substituting (4) in (3), we get

$$\beta_1 + \beta_2 + \beta_3 + \cdots + \beta_{s+1} \quad = \quad n - i - s + 1 \tag{5}$$

Equations (4) and (5) are Diophantine equations and they have solutions $\binom{i+s-1}{s-1}$ and $\binom{n-i-s}{s}$, respectively. Therefore, the total number of solutions is $\binom{i+s-1}{s-1} \times \binom{n-i-s}{s}$. Now for a given combination (α_1, α_2, $\cdots$, α_s, β_1, β_2, $\cdots$, β_{s+1}), the probability that the combination will take place is:

$$p^{\alpha_1+\alpha_2+\alpha_3+\cdots+\alpha_s} \times (1-p)^{n-(\alpha_1+\alpha_2+\alpha_3+\cdots+\alpha_s)}$$

The probability of Case-(i) is then given by:

$$P_1 \quad = \quad \sum_{s=1}^{\lceil \frac{n}{2} \rceil - 1} \binom{i+s-1}{s-1} \times \binom{n-i-s-1}{s} \times p^{i+s} \times (1-p)^{n-i-s}$$

Case(ii): $\{\beta_{s+1} = 0$ **and** $\beta_1 \geq 1\}$: Now in the last block all the α_s packets will be lost. So,

$$\begin{aligned} (\alpha_1 - 1) + (\alpha_2 - 1) + \cdots + (\alpha_s - 1) + \alpha_s &= i \quad (or) \\ \alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_s &= i + s - 1 \end{aligned} \tag{6}$$

The total number of solutions for the above equation is $\binom{i+s-2}{s-1}$. Substituting (6) in (3) gives:

$$\beta_1 + \beta_2 + \beta_3 + \cdots + \beta_s \quad = \quad n - i - s + 1 \tag{7}$$

The total number of solutions for the above equation is $\binom{n-i-s}{s-1}$. Therefore, the total number of solutions is $\binom{i+s-2}{s-1} \times \binom{n-i-s}{s-1}$. Now for a given combination (α_1, α_2, $\cdots$, α_s, β_1, β_2, $\cdots$, β_{s+1}), the probability that the combination will take place is:

$$p^{\alpha_1+\alpha_2+\alpha_3+\cdots+\alpha_s} \times (1-p)^{n-(\alpha_1+\alpha_2+\alpha_3+\cdots+\alpha_s)} \tag{8}$$

The probability of Case-(ii) is then given by:

$$P_2 \quad = \quad \sum_{s=1}^{\lfloor \frac{n}{2} \rfloor} \binom{i+s-2}{s-1} \times \binom{n-i-s}{s-1} \times p^{i+s-1} \times (1-p)^{n-i-s+1}$$

Case(iii): $\{\beta_{s+1} \geq 1$ **and** $\beta_1 = 0\}$: Now in the last block all the α_s packets will not be lost. So,

$$\begin{aligned} (\alpha_1 - 1) + (\alpha_2 - 1) + \cdots + (\alpha_s - 1) + (\alpha_s - 1) &= i \qquad (or) \\ \alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_s &= i + s \end{aligned} \tag{9}$$

The total number of solutions for the above equation is $\binom{i+s-1}{s-1}$. Substituting (9) in (3) gives:

$$\beta_2 + \beta_3 + \cdots + \beta_s + \beta_{s+1} \quad = \quad n - i - s \tag{10}$$

The total number of solutions for the above equation is $\binom{n-i-s-1}{s-1}$. Therefore, the total number of solutions is $\binom{i+s-1}{s-1} \times \binom{n-i-s-1}{s-1}$. Now for a given combination (α_1, α_2, $\cdots$, α_s, β_1, β_2, $\cdots$, β_{s+1}), the probability that the combination will take place is:

$$p^{\alpha_1+\alpha_2+\alpha_3+\cdots+\alpha_s} \times (1-p)^{n-(\alpha_1+\alpha_2+\alpha_3+\cdots+\alpha_s)} \tag{11}$$

The probability of Case-(iii) is then given by:

$$P_3 \quad = \quad \sum_{s=1}^{\lfloor \frac{n}{2} \rfloor} \binom{i+s-1}{s-1} \times \binom{n-i-s-1}{s-1} \times p^{i+s} \times (1-p)^{n-i-s}$$

Case(iv): $\{\beta_{s+1} = 0$ **and** $\beta_1 = 0\}$: Now in the last block all the α_s packets will be lost. So,

$$\begin{aligned} (\alpha_1 - 1) + (\alpha_2 - 1) + \cdots + (\alpha_{s-1} - 1) + \alpha_s &= i \;\; (or) \\ \alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_s &= i + s - 1 \end{aligned} \tag{12}$$

The total number of solutions for the above equation is $\binom{i+s-2}{s-1}$. Substituting (12) in (3) gives:

$$\beta_2 + \beta_3 + \cdots + \beta_s \quad = \quad n - i - s + 1 \tag{13}$$

The total number of solutions for the above equation is $\binom{n-i-s}{s-2}$. Therefore, the total number of solutions is $\binom{i+s-2}{s-1} \times \binom{n-i-s}{s-2}$. Now for a given combination (α_1, α_2, $\cdots$, α_s, β_1, β_2, $\cdots$, β_{s+1}), the probability that the combination will take place is:

$$p^{\alpha_1+\alpha_2+\alpha_3+\cdots+\alpha_s} \times (1-p)^{n-(\alpha_1+\alpha_2+\alpha_3+\cdots+\alpha_s)} \tag{14}$$

The probability of Case-(iv) is then given by:

$$P_4 = \sum_{s=2}^{\lceil \frac{n}{2} \rceil} \binom{i+s-2}{s-1} \times \binom{n-i-s}{s-2} \times p^{i+s-1} \times (1-p)^{n-i-s+1}$$

Corner Cases: The following two cases are not covered by the above cases, *i.e.,* (i) Number of βs is 0 and (ii) Number of αs is 0 : whose probabilities are p^n (the probability for all the packets to be lost) and $(1-p)^n$ (the probability of no packet to be lost), respectively.

So, the probability of losing *i* packets out of *n* packets at the receiver after applying an *FEC-Offset* of 1 is:

$$\begin{aligned} p_1(i,n) \quad &= \quad P_1 + P_2 + P_3 + P_4 + (p^n) \quad (if \;\; i \; = \; n) \\ &+ \;\; (1-p)^n \;\; (if \; i \; = \; 0) \end{aligned} \tag{15}$$

Hence the *residual packet loss rate, RPLR*, at the receiver node is:

$$RPLR(1, P_{eloss}) \quad = \quad \sum_{i=1}^{n} \frac{i \times p_1(i,n)}{n} \tag{16}$$

B. Calculation of *RPLR* when *FEC-Offset* = *r*

Fig. 6 shows the way of protecting packets with *FEC-Offset* = *r*. Let us arrange the total number of packets into *m* disjoint sets as shown in Fig. 7. The sequence for *k* = *1*, *2*, *3*, *4*, $\cdots$, *r* will cover all the packets. The total number of disjoint sets can be calculated as follows:

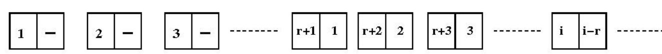

Fig. 6. **Packet transmission with** ***FEC-Offset*** **=** ***r*****.**

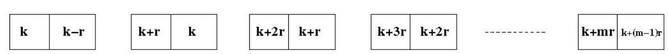

Fig. 7. **Dividing total packets into** m **disjoint sets.**

$$\begin{aligned} k + mr \quad &\leq \quad n \\ m \quad &\leq \quad \frac{n-k}{r} \\ m_{max} \quad &= \quad \left\lfloor \frac{n-k}{r} \right\rfloor \\ 0 \quad &\leq \quad m \quad \leq \quad \left\lfloor \frac{n-k}{r} \right\rfloor \end{aligned}$$

$$\text{So, totally there are} \quad \left\lfloor \tfrac{n-k}{r} \right\rfloor + 1 \quad \text{disjoint sets.}$$

Loss of *i* packets can be written as losing (see Fig. 7)

$$\begin{aligned} i_1 \;\; \text{packets with a probability} \;\; p_1(\left\lfloor \frac{n-1}{r} + 1 \right\rfloor, i_1) \; for \; k \quad &= \quad 1 \\ i_2 \;\; \text{packets with a probability} \;\; p_1(\left\lfloor \frac{n-2}{r} + 1 \right\rfloor, i_2) \; for \; k \quad &= \quad 2 \\ . \quad &\quad \;\; . \\ i_r \;\; \text{packets with a probability} \;\; p_1(\left\lfloor \frac{n-r}{r} + 1 \right\rfloor, i_r) \; for \; k \quad &= \quad r \\ \text{such that} \;\; i = i_1 + i_2 + \cdots + i_r & \end{aligned}$$

So, the probability of losing i packets out of n packets in case of *FEC-Offset* = r is given by:

$$p_r(i,n) = \prod_{i_1}\prod_{i_2}\cdots\prod_{i_r} p_1(\left\lfloor \frac{n-1}{r}+1 \right\rfloor, i_1) p_1(\left\lfloor \frac{n-2}{r}+1 \right\rfloor, i_2) \cdots p_1(\left\lfloor \frac{n-r}{r}+1 \right\rfloor, i_r) \quad (17)$$

Hence the *residual packet loss rate, RPLR*, at the receiver node is:

$$RPLR(r, P_{eloss}) = \sum_{i=1}^{n} \frac{i \times p_r(i,n)}{n} \quad (18)$$

C. Validation

In this subsection, we validate our Packetization scheme. We use the simulation parameters that are shown in Table I. The destination node calculates P_{eloss} for every 1000 packets and sends back to the sender node. The source node then estimates the *RPLR* (using Eqn. (18) for different r values and finds the best *FEC-Offset*, r and use it for subsequent packet transmissions) and the destination node calculates the *RPLR* based on the number of packets. Fig. 8 shows the change in frame loss rate for average bad time (average time spent in bad state, see next section). As observed in Fig. 8 the simulation values closely match with analytical results.

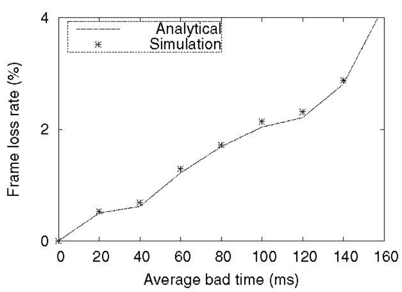

Fig. 8. **Variation of frame loss rate vs. average bad time.**

IV. PERFORMANCE EVALUATION

We simulated our scheme using NS-2 Network Simulator [7]. The simulation parameters are shown in Table 1. For all cases of mobility in the network, we set the pause time to 0 *sec* and also set the minimum and maximum speeds to the same value to ensure that the nodes move at a constant speed. The RTS/CTS mechanism is disabled in our simulations since all the packets that we consider are small sized packets. Simulation runs are carried out for 20 seeds averaged over 10 flows and all the results conform to 95% confidence levels. We modified the Dynamic Source Routing (DSR) [8] protocol to obtain two maximally node-disjoint routes from source to destination. We simulated bursty packet losses using the *Gilbert-Elliott model* provided in NS-2. The average dwell time in the good state is 1000 *ms*. While varying mobility, the average bad time (the average time spent in the bad state) is kept constant at 30 *ms*.

A. Packetization scheme vs. Layered scheme vs. MD scheme

We compare the following two schemes with our scheme to evaluate the effectiveness of our proposed Packetization scheme. In all the schemes, the 802.11 DCF with 0-MAC retransmissions is used to ensure fairness while comparing with one another.

TABLE I
SIMULATION PARAMETERS

Parameter	Value	Parameter	Value
Terrain Area	1000 *m* x 1000 *m*	Transmission Range	250 *m*
Channel Capacity	11 Mbps	# of Nodes	75
Mobility Model	*Random Way Point* (RWP)	Simulation Duration	300 *s*
Background Traffic Arrival Rate	30 *pkts/s*	# of Background Traffic Flows	10
# of Frames per Voice Flow	15000	# of Voice Flows	10
MAC Protocol	802.11 DCF	Inter Frame Time	20 *ms*
Traffic Type	CBR	Voice Frame Size	32 B

TABLE II
PACKET OVERHEADS IN VARIOUS SCHEMES

Parameter	Layered Scheme (bytes)	MD Scheme (bytes) [1]	Packetization Scheme (bytes)
Path 1 Payload Size	34	46	28/36
Path 2 Payload Size	48	46	28/36
Total Payload Size	34+48=82	46+46=92	36+36=72
Overhead due to (UDP+IP+802.11 MAC) for 2 paths	2 * (8 + 20 + 24)=104	2 * (8 + 20 + 24)=104	2 * (8 + 20 + 24)=104
Total Size	186	196	176

- **Layered Scheme** : In this scheme, the raw voice stream is encoded into the BL sub-stream and one EL sub-stream using a layered codec. Multipath transport is used to transmit the BL sub-stream along one path and the EL sub-stream along another path. Both the BL and EL sub-streams carry a FEC copy of previous BL sub-stream.
- **MD Scheme** : In this scheme, using MD coding the raw voice stream is encoded into two descriptions [3]. Both the descriptions are protected using FEC with offset 1.

For all the simulations the $delay_{threshold}$= 200 *ms*. In the Packetization scheme for simulation purposes, the destination node periodically (every 1000 packets) calculates the network loss parameter, P_{eloss} and feeds back to the source node using an explicit UDP packet (similar to RTCP receiver reports via UDP). As shown in Table II, our scheme has 10.2 % and 5.57 % less overhead compared to MD and Layered schemes, respectively. The following subsections discuss the performance issues of these schemes.

1) **Effect of Average Bad Time and Mobility:** Fig. 9 shows the *frame loss rate* for various schemes by varying Average Bad Time (ABT). Since, Packetization scheme adaptively adjusts its *FEC Offset* based on the network state, it is able to sustain burst losses better compared to other two schemes. The variation of *average size of bursts* for varying ABT is given in Fig. 10. When the channel gets bursty, the average burst size of Layered scheme gets worsened and it increases rapidly which indicates that most of the BL packets are lost either due to channel errors or due to collisions. On the other hand the Packetization scheme has better performance compared to other schemes due to its adaptive FEC. Similar behavior can be seen in Fig. 11 for varying mobility.

2) **Measurement of Perceptual Evaluation Speech Quality Mean Opinion Score (PESQ-MOS):** At the destination, the voice frames are decoded and the wide band version of

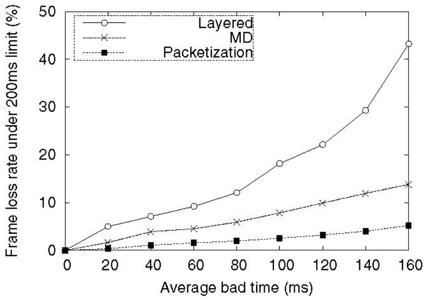

Fig. 9. **Variation of frame loss rate vs. average bad time under 200 *ms* threshold.**

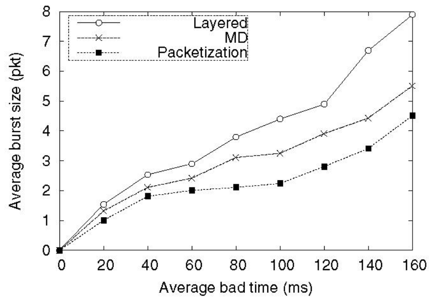

Fig. 10. **Variation of average burst size vs. average bad time.**

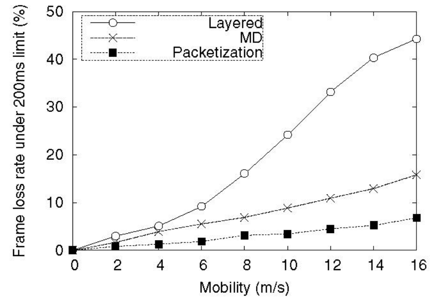

Fig. 11. **Variation of frame loss rate vs. mobility under 200 *ms***

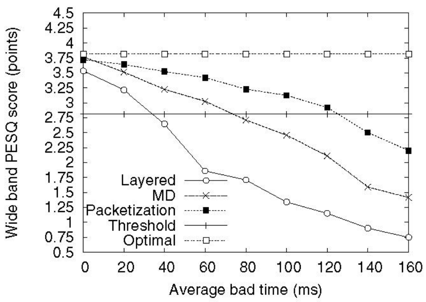

Fig. 12. **Variation of wide band PESQ score vs. average bad time.**

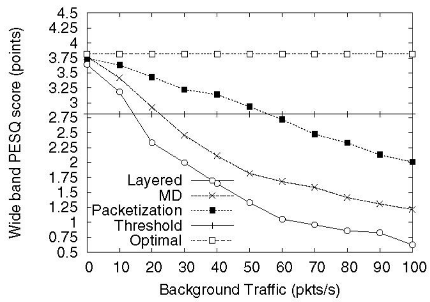

Fig. 13. **Variation of wide band PESQ score vs. background traffic.**

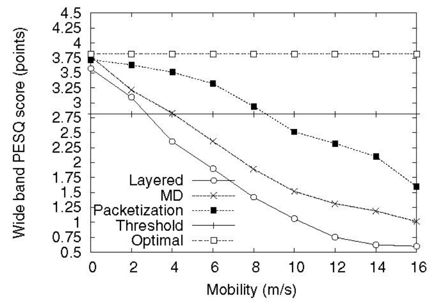

Fig. 14. **Variation of wide band PESQ score vs. mobility.**

ITU perceptual measurement algorithm, PESQ Mean Opinion Score (MOS) reference software tool [9] is used to measure their *perceived voice quality*. The PESQ compares the degraded speech with the reference speech and computes the objective MOS value in a 5-point score ranging from -0.5 (worst) to 4.5 (best). With respect to a original raw voice frame, the voice quality scores of different voice frames are evaluated using PESQ-MOS reference software tool. The evaluated voice quality scores of (a) raw voice frame, (b) decoded AMR-WB voice frame, (c) decoded BL of AMR-WB voice frame of Packetization scheme, (d) decoded BL+MDC1/MDC2 of EL of AMR-WB voice frame of Packetization scheme, (e) decoded description of AMR-WB voice frame of MD scheme are 4.5 (Ideal Quality), 3.818 (Optimal), 2.81 (Basic), 3.24, and 2.86, respectively. The optimal quality score (3.818) corresponds to the decoding of AMR-WB (lossy encoded) voice frame assuming no losses in the network. The threshold quality score corresponds to decoding of BL of AMR-WB voice frame. Both Optimal and threshold scores are shown in Figures 12, 13, and 14. Fig. 12, Fig. 13, and Fig. 14 show the measured wide band PESQ-MOS score at the destination for varying ABT, background traffic and mobility respectively. As observed in the graphs, the Packetization scheme exhibits better quality compared to both Layered and MD schemes. The increase in performance gain in Packetization scheme is due to the reason 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*. Thus, our scheme provides the best quality with the minimal overhead at all the times compared to the existing schemes.

V. 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 ad hoc wireless networks. We measured perceived speech quality for different schemes by integrating NS-2 Network Simulator with a real adaptive speech codec (AMR-WB codec) and a perceived quality evaluation system based on the wide band version of ITU-T PESQ. We provided an analytical approach to estimate the *residual packet loss rate*. The simulation results showed that the proposed Packetization scheme outperforms the existing schemes in terms of *frame loss rate*, *end-to-end delay*, and wide band PESQ-MOS.

REFERENCES

[1] Chi-hsien Lin, Hui Dong, Upamanyu Madhow, and Allen Gersho, "Supporting Real-time Speech on Wireless Ad Hoc Networks: Inter-packet Redundancy, Path Diversity, and Multiple Description Coding," *in Proceedings of ACM Workshop on WMASH*, pp. 11–20, October 2004.

[2] ITU-T Recommendation, G.114: "One Way Transmission Time," February 1996.

[3] H. Dong, A. Gersho, J. Gibson, and V. Cuperman, "A Multiple Description Speech Coder Based on AMR-WB for Mobile Ad Hoc Networks," *in Proceedings of IEEE ICASSP*, vol. 1, pp. 277–280, May 2004.

[4] Shiwen Mao, Shunan Lin, S. S. Panwar, Yao Wang, and E. Celebi, "Video Transport over Ad hoc Networks: Multistream coding with Multipath transport," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 10, pp. 1721–1737, December 2003.

[5] J-C. Bolot, S. Fosse-Parisis, and D. Towsley, "Adaptive FEC-Based Error Control for Internet Telephony," *in Proceedings of IEEE INFOCOM*, vol. 3, pp. 1453–1460, March 1999.

[6] 3GPP TS 26.171: Adaptive Multi-Rate Wide Band Speech Codec, http://www.3gpp.org/ftp/Specs/html-info/26171.htm.

[7] The Network Simulator - NS-2 http://www.isi.edu/nsnam/ns.

[8] D. B. Johnson and D. A. Maltz, "Dynamic Source Routing in Ad hoc Wireless Networks," *Mobile Computing*, Kluwer Academic Publishers, vol. 353, pp. 153–181, 1996.

[9] 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.