Multiple Descriptions and Path Diversity Using the AMR-WB Speech Codec for Voice Communication over MANETs

Jagadeesh Balam
Department of Electrical and Computer Engineering
University of California, Santa Barbara, 93106.
balam@ece.ucsb.edu

Jerry D. Gibson
Department of Electrical and Computer Engineering
University of California, Santa Barbara, 93106.
gibson@ece.ucsb.edu

ABSTRACT

We compare different source diversity methods for converstional voice communication over multiple routes in a mobile ad-hoc network (MANET). A new multiple description (MD) codec based on the AMR-WB codec, with two balanced side descriptions (6.9 kbps each) is presented. We compare the performance of the MD codec against two other diversity methods, 1) duplicating speech encoded with AMR-WB at 6.6 kbps and 2) duplicating speech encoded with AMR-WB at 12.65 kbps. We show that because of the large packet headers added to each packet by typical MANET protocols, the overhead of sending the simple path diversity methods is not much larger than the overhead for sending MD streams over different paths, and the gain in speech quality we get from duplicating AMR-WB at 12.65 kbps over sending MD codec streams is significant. We compare the speech quality delivered by each of the methods under random and bursty packet loss conditions. The quality of decoded speech is evaluated using WPESQ, a wideband extension to the PESQ algorithm.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: [wireless communication]

General Terms

Reliability, Design, Performance

Keywords

Voice communications, voice quality indicator, AMR-WB, 802.11, WPESQ, MOS, multiple descriptions, path diversity, mobile ad-hoc networks

IWCMC'06, July 3–6, 2006, Vancouver, British Columbia, Canada.

1. INTRODUCTION

Mobile Ad-hoc Networks (MANETs) are formed by mobile wireless hosts without the need of an existing infrastructure. MANETs are seen as future networks for military environments, emergency operations and office or conference environments. Conversational voice communication over a wireless MANET is a challenging problem because of the error prone wireless channel, the changing topology of the network, delays involved in establishing a new link or finding a new route, and the current MAC protocols which were not developed for real-time communication. Ad-hoc networks based on the IEEE 802.11 standard is an example. The IEEE 802.11 standard is designed primarily for non-real time transfer of data and these protocols may not be suitable for conversational voice communication. IEEE 802.11 MAC protocols are designed to minimize collisions and depend on retransmissions to ensure successful transmission of a packet irrespective of the delay incurred by the packet. Interactive communication cannot tolerate large delays (larger than 200 ms according to ITU-T Recommendation G.114) and in multi-hop communications unknown delays will occur at each intermediate node in the route.

The widespread availability of 802.11 based WLANs and the possibility of supporting low-cost wireless voice communications has attracted significant interest in developing solutions for reliable voice communications over these networks. Most of the efforts have been toward adapting the 802.11 MAC layer for reducing retransmissions and packet losses. Changes to the MAC layer are suggested in [1, 2, 3] so that voice packets are dropped only when there are errors in perceptually important bits. Such schemes accept some packets with errors in voice payloads and decode the corresponding speech frames as they are, instead of considering the packets as lost and using a frame error concealment algorithm to conceal the whole frames. These methods lead to a reduced number of retransmissions, smaller end-to-end delay, and less traffic in the network. The quality of delivered speech is observed to be significantly improved [3]. Commercial WLAN phones by *Spectralink* [4] use a priority scheme for transmission of speech packets to minimize delay for these packets. The link layer also uses zero back-off for speech packets instead of the exponential random back-off scheme usually used for data transfer. Forward error correction for only perceptually important packets is suggested in [5] and [6]. In [5], perceptually important packets are determined by computing an analysis-by-synthesis distortion for different parameters in an encoded voice frame. The first fifteen frames after an unvoiced/voiced transition are classified as important packets and these frames are protected by both increasing the maximum number of retransmissions and by sending additional copies in [6]. Servetti and Martin [7] suggest using a variable bit-rate codec to adapt the bit rate of the speech encoder according to instantaneous channel conditions.

An important method to improve reliability of transmission over a MANET is to use path diversity, i.e. send data simultaneously through multiple paths. Each packet can be repeated over independent paths and since the probability of all the paths breaking down simultaneously is smaller than a single channel breaking down, the probability of packet loss is reduced. However, sending multiple copies of the same packet is inefficient usage of bandwidth. To improve bandwidth efficiency, a source coding diversity method like multiple description (MD) coding can be used. In MD coding, multiple descriptions/bit-streams of the source are created in such a way that each description can be used to reconstruct the source with acceptable quality and two or more descriptions can be combined to give a better quality reconstruction.

We designed a new balanced MD coder based on the AMR-WB [8] codec, where the two descriptions are of the same rate and similar quality. We compare the performance of MD coding with respect to the simple path diversity methods of 1) repeating a single description (SD) codec with a bit rate about the same as each description of the MD codec (AMR-WB @ 6.6 kbps) and 2) repeating the SD codec (AMR-WB @ 12.65 kbps) over the independent paths. Although these path diversity methods seem inefficient, if packet headers are taken into consideration, we see that these methods are comparable to the MD method in terms of bandwidth efficiency. We also see that the second path diversity method provides consistently good speech quality compared to the MD method. Performance is measured in terms of quality of speech delivered to the receiver and is evaluated using WPESQ, a wideband extension to PESQ.

2. MD-AMR: A MULTIPLE DESCRIPTION SPEECH CODER BASED ON AMR-WB SPEECH CODEC

The Adaptive Multirate Wideband (AMR-WB) [8] speech codec was selected in December 2000 for GSM and the third generation mobile communication WCDMA system for providing wideband speech services. It was also selected as recommendation G.722.2 by the ITU-T. AMR-WB operates on speech of extended bandwidth ranging from 50Hz to 7000Hz. Traditionally, speech codecs were designed for narrowband speech of telephone bandwidth (200 to 3400 Hz), but the evolution of broadband multimedia services has spawned an increased interest in wideband speech. Wideband speech sounds more natural and is more intelligible than the traditional narrowband speech. AMR-WB is particularly popular because of it's accepatance by the 3GPP for GSM and WCDMA and also ITU-T for wireline wideband speech services. The AMR-WB speech codec is an ACELP (Algebraic Codebook Excited Linear Prediction) based codec and operates on 20ms frames. The AMR-WB codec operates in nine different modes with bit-rates ranging from 6.6 to 23.85 kbps. Mode 2 operating at 12.65 kbps and other modes above it offer high quality speech. We use AMR-WB modes 0 (6.6 kbps) and 2 (12.65 kbps) in our experiments for comparison of different path diversity methods.

We designed an MD coder using the AMR-WB codec in mode 2 (12.65 kbps) (MD-AMR), based on the MD codec introduced in [9]. Our MD coder creates balanced descriptions, i.e. each description is of the same rate, and speech decoded from either descriptions is of similar quality. Such a codec is more suitable for an ad-hoc network, because in a MANET, we cannot guarantee delivery or a better QoS for any one path. The idea behind the coder is to take an SD coder (AMR-WB) and split the bit stream into two substreams. This is similar to the no-excess joint rate case of MD coding, where the individual descriptions can be combined to give an optimal joint description. Since dividing the bit stream into two non-overlapping portions cannot give us acceptable quality at the side decoders, we inject some redundancy by replicating vital information in both the descriptions. The distortion at the central decoder is still the same as the SD decoder but the effective bit-rate is higher due to the redundancy introduced in the side descriptions. Of course, the quality delivered by each description will be worse than that of an SD codec optimized for the same rate as an individual description.

2.1 MD-AMR Encoder

The encoder of the MD-AMR coder (Fig. 1) consists of the AMR-WB encoder and a bit stream splitting block that divides the AMR-WB encoded bit stream into two different bit streams in such a way that both the descriptions get about the same amount of information. The AMR-WB encoder divides the speech frame into four sub-frames for estimating various parameters required for CELP coding. Except for the LPC coefficients, all the other parameters are determined for each sub-frame. The simplest way to divide the bit stream would be on a subframe basis. We can send the bits corresponding to different sub-frames as different descriptions on independent channels/paths. When all the descriptions are received, they can be combined to reconstruct the whole speech frame. When only a subset of descriptions is received, the subframes corresponding to the received descriptions can be reconstructed and the missing subframes can be concealed with the information from the nearest received sub-frames.

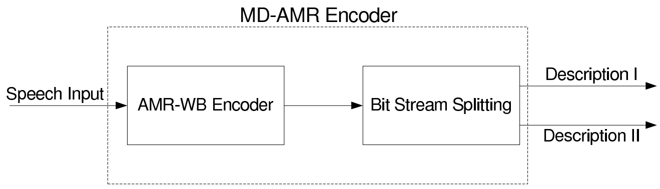

Figure 1: Encoder of the MD-AMR coder

The bit splitting for various parameters in the MD-AMR encoder is described below:

2.1.1 *ISF*

Although the LPC parameters are computed only once per frame, they are encoded (after conversion to Immitance

Spectral Frequencies (ISFs)) using split-multistage vector quantization. The structure of the vector quantizer gives us the flexbility to divided the bits corresponding to the LPC coefficients in such a way that they can be decoded to give at least a coarse reproduction of the parameters. The multistage quantization allows us to have the important first layer to be repeated in both the descriptions, so that either description has at least a coarse reproduction of the LPC coefficients.

The ISFs are coded using a 2 stage split vector quantizer. Sixteen bits are used to code the index of the the code vectors in the first stage. These 16 bits are included in both the descriptions, because without this information LPC parameters cannot be reconstructed at all. The next stage vector is divided into 5 sub-vectors and the five sub-vectors are coded with $6 + 7 + 7 + 5 + 5 = 30$ bits. The bits corresponding to the first sub-vector are included in both the descriptions, while the bits corresponding to the second and the fifth sub-vectors are included in description I only and the remaining bits corresponding to the third and fourth sub-vectors are included in description II only. This way of splitting was experimentally determined to give the most symmetric quality at the two descriptions.

2.1.2 Pitch Delay for Adaptive codebook

Pitch delay is calculated on a sub-frame basis, but the second sub-frame pitch delay is differentially encoded with respect to the first sub-frame. Without the first (third) sub-frame pitch delay, the second (fourth) sub-frame bits are useless, so the first and second sub-frame bits are included in description I and the third and fourth sub-frame bits are included in description II.

2.1.3 Adaptive and Fixed codebook gains

In AMR-WB, gains for Adaptive and fixed codebook are jointly quantized with seven bits in each sub-frame. We include bits corresponding to the first and the third sub-frame gains in description I and the second and the fourth sub-frame information in description II. This is done because when only one description is received, the missing gain information is concealed using the previous sub-frame information. If only description I is received, second sub-frame gains are concealed using the first sub-frame gains and fourth sub-frame gains are concealed using third sub-frame information. If instead of allocating alternate sub-frames to each description, we include first and second sub-frame in description I and third and fourth sub-frame in description II, then, if only description I is received, fourth sub-frame gains can be concealed only using with the second sub-frame gains. This is worse than concealing fourth sub-frame gains using information from the third sub-frame, which is the closest sub-frame.

2.1.4 Fixed codebook Indices

AMR-WB uses 36 bits per sub-frame to code the fixed codebook. Here again, first and third sub-frames are included in description I and second and fourth sub-frames are included in description II.

Table 1 shows the bit allocation for the two descriptions. The table as a whole shows the bit allocation for each parameter in the bit stream of AMR-WB, mode-2 (12.65 kbps). The numbers within the parentheses indicate that the corresponding bits belong only to description II and the bits corresponding to the emphasized numbers are replicated in both the descriptions. The remaining bits belong only to description I. The VAD flag is included in both the descriptions.

Table 1: Bit allocation for the MD codec based on AMR-WB

ISP	Stage 1: ***8 8***				I,(II)
	Stage 2: ***6*** 7 (7) (5) 5				34,(34)
	1st sf	2nd sf	3rd sf	4th sf	
VAD					1,(1)
LTP-filtering	1	1	(1)	(1)	2,(2)
Pitch delay	9	6	(9)	(6)	15,(15)
Algebraic Code	36	(36)	36	(36)	72,(72)
Gains	7	(7)	7	(7)	14,(14)
Total					138,(138)

The bit-rate for each description is 6.9 kbps and the bit-rate for the combination is 13.8 kbps of which 1.15 kbps is redundant. The redundant bit-rate is the penalty paid to make the distortion at the side decoders acceptable. Each description sounds worse than AMR-WB at 6.6 kbps since they are obtained by splitting the rate of a higher rate codec compared to AMR-WB@6.6 kbps which is optimized to give the best quality at that rate.

2.2 Decoder

Figure 2 shows a block diagram of the MD-AMR decoder. When both the descriptions created using the MD-AMR encoder are delivered at the receiver, the bit streams are combined to form the AMR-WB bit stream and the AMR-WB decoder is used to reconstruct the signal. When both the descriptions are lost, AMR-WB packet loss concealment is used to conceal the lost packet. When only one description is received, the missing bits when compared with the AMR-WB bit stream are substituted using information from the most recent frame received in the MD-AMR decoder.

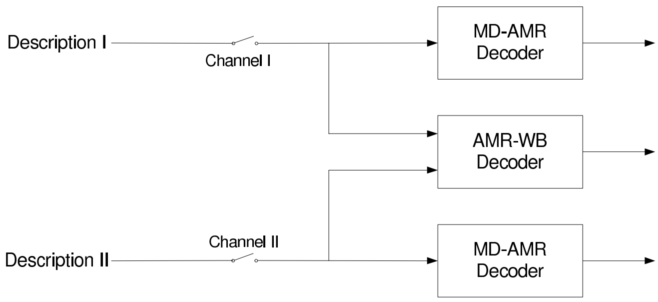

Figure 2: Decoder of the MD-AMR coder

The decoding process of either description is similar. In the following points, we summarize the decoding of description I when description II is lost:

- The sub-vectors corresponding to the missing bits in the ISP indices are ignored and not added in the second stage of the vector quantizer
- The pitch-lag values of the third and fourth subframes are substituted with the pitch-lag value of the second subframe

- LTP-filtering flag of the third and fourth subframe is set to be the same as that of 2nd subframe
- The fixed code vector second (fourth) subframe is set to be the same as that of the first (third) subframe and the gains for the second (fourth) subframe are set to the gains of first (third) subframe attenuated by 3dB

3. EXPERIMENTS

The objective of our experiments is to compare the end quality of speech delivered by three different diversity methods for communication over two independent paths and also show the effect of packet headers on their performance. We call AMR-WB at 12.65 kbps the full-rate (FR) codec and AMR-WB @6.6 kbps the half-rate (HR) codec. The three diversity methods we consider are, 1) sending the two descriptions of the MD codec over the two paths - Multiple Descriptions with Path Diversity (MD-PD), 2) duplicating the full rate codec over the two independent paths (DFR-PD) and 3) duplicating the half-rate codec over the two independent paths (DHR-PD). We also consider the SD case where the full-rate codec is sent over a single channel without any path diversity. Sizes of encoded frames for different codecs are listed in Table 2.

Table 2: Frame sizes for the codecs

Codec	Frame size (bits)
AMR@12.65kbps (FR)	253
AMR@6.6 kbps (HR)	132
MD	138

3.1 Setup

We assume that two independent paths with similar channel conditions are available between the sender and the receiver. We follow the 802.11 concept wherein a speech packet is dropped if even one of the bits in the packet is in error. We do not use retransmissions when packets are lost in the network, but rely on path diversity and packet loss concealment to overcome packet losses due to bit errors in the channel. We consider two kinds of packet losses, 1) random packet losses due to random bit errors in the channel and 2) bursty packet losses due to phenomena like fading or shadowing in the network or other factors like a link failure. In the case of random bit errors, for a given bit error rate (BER), we first find packet loss probability p using Eq. (1), for the packet length of each codec,

$$p = 1 - (1 - BER)^L \tag{1}$$

where L is the length of the packet in bits. For random errors, the packet headers are a concern because they increase the length of the packet and hence increase p. For each p, frames were dropped randomly in the encoded speech files using 250 different seeds for the random generator. We used six different (3 male, 3 female) speech files in our experiments. Each file is about 8 seconds long and consists of two different sentences spoken by the same speaker. For all our experiments, we assume that each packet contains one 20 ms speech frame.

Quality of the decoded speech was evaluated using WPESQ (Wideband PESQ). WPESQ is an extension to ITU-T recommendation P.862 [10], proposed in [11], to adapt PESQ (Perceptual Evaluation of Speech Quality) for use in measuring wideband speech quality. The difference between WPESQ and PESQ is only the input filter characteristics, the psychoacoustic model and the error model are however the same. We use an implementation based on this proposal to evaluate decoded speech quality in our experiments. WPESQ scores have been observed to be different from the actual subjective MOS scores. To improve the comparability of WPESQ scores with subjective MOS values, Barriac et al. propose a function for mapping WPESQ scores to subjective MOS values in [12]. This function is given as

$$y = 1 + \frac{4}{1 + e^{-2x+6}} \tag{2}$$

where x is the WPESQ score and y is the corresponding mapped value. We use the above function to convert the WPESQ scores to WPESQ-MOS values.

3.2 Random packet losses

3.2.1 No Packet Headers

First, we look at a scenario where there are no packet headers added to the speech packets. The channel has random bit errors and the packet loss probability is p given by Eq. (1). Table 2 lists the packet sizes for the different codecs used. SD and DFR-PD have the largest p for a given BER, because they use the FR codec with highest frame size, while MD-PD and DHR-PD have almost equal p because the difference in their packet sizes is very small. For each BER, the average mapped WPESQ scores (WPESQ LQ MOS) are plotted in Fig. 3. We can see that at the lowest BER of 10^{-5}, there is already a loss of performance in SD. Note that without any packet losses SD, MD-PD and DFR-PD should give the same MOS values. At a BER of 10^{-5}, there is no loss in the performance of DFR-PD as at least one path successfully transmits all the time. MD-PD is also not affected at this BER because of the small packet size and a corresponding very small packet loss rate ($\approx .14\%$). If we consider SD, MD-PD and DHR-PD as requiring the same bandwidth for transmission, then MD-PD is a clear winner except for very high BERs, when DHR-PD starts doing better than MD-PD. This is because, at such high BERs only one of the paths is up for most of the time and a single description of the MD codec sounds worse than AMR@6.6 kbps. DFR-PD delivers the best quality of speech but at a penalty of required bandwidth for transmission. The number of bits required to be sent for DFR-PD is almost double that either MD-PD or DHR-PD.

3.2.2 Packet Headers

Now we consider a more realistic scenario where headers are added to the speech packets by the lower protocol layers. In a typical 802.11 based ad-hoc network, headers would be added by RTP, UDP, IP and the 802.11 MAC layer protocol. The overheads for each packet add up to 68 bytes (the 802.11 MAC (28 bytes), IP (20 bytes), UDP (8 bytes) and RTP (12 bytes)), significantly larger than the payload, which is a maximum of 32 bytes in our experiments. Table 3 shows the effective packet sizes of each codec after the inclusion of packet headers. Payloads are padded with zero bits to form complete octets.

The performance of all the methods with these new packet sizes for changing BERs is shown in Fig. 4. The large advan-

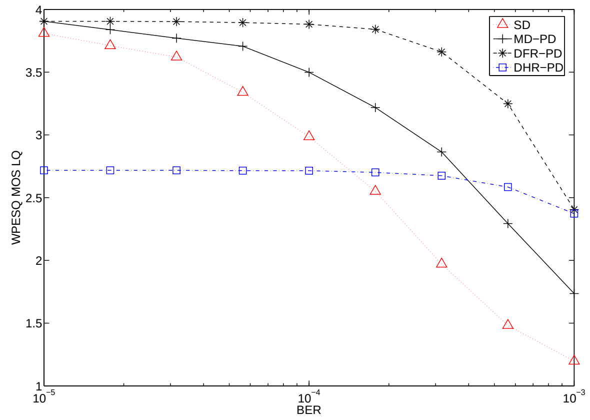

Figure 3: Average WPESQ-MOS values, for changing BER, without any packet headers

Table 3: Packet sizes with headers

Codec	Packet size (bytes)
AMR@12.65kbps (FR)	100(32+68)
AMR@6.6 kbps (HR)	85(17+68)
MD	86(18+68)

tage obtained with DFR-PD over MD-PD (WPESQ MOS gain of 1.71 at 10^{-4} BER) requires only a 16% increase in the number of bits transmitted per path. Observe from Figs. 3 and 4 that the overall performance of all the methods drops because of the increased packet loss rates (PLR) resulting from the larger packet sizes. The ordering of the codecs with respect to their performance is still maintained, but MOS for MD-PD drops below that of DHR-PD at a smaller BER again because of the higher PLRs and only one path being up for most of the time. Also, compared to the case with no headers (Fig. 3), the advantage of MD-PD over SD narrows considerably.

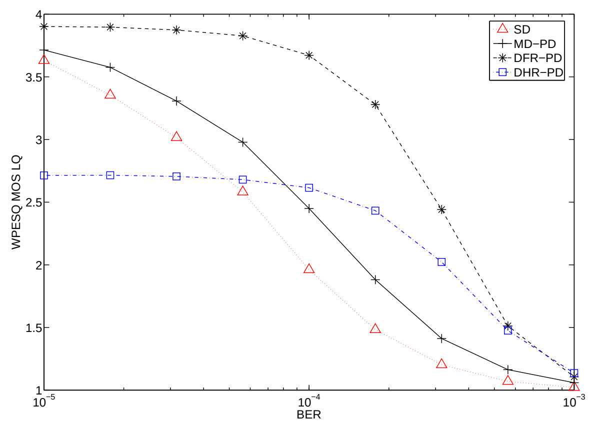

Figure 4: Average WPESQ-MOS values, for changing BER, with packet headers

Using larger payloads by coding longer frame sizes or including more frames per packet might reduce the inefficiency due to the headers, but doing so also increases the latency, which is a principal concern in conversational voice communication. The best possible solution for the problem of large packet headers today is using a header compression scheme like RoHC (Robust Header Compression). Efforts are underway to make RoHC compatible with IEEE 802.11. Using RoHC the IP/UDP/RTP headers can be compressed to very small sizes of up to one byte. If we assume an average compressed header size of 2 bytes, the MAC layer header is still of significant size (28 bytes), and the ratio of MD to FR packet sizes is still around 0.77. Figure 5 shows the performance of all the methods with reduced packet sizes because of compressed headers. DFR-PD still performs significantly better than MD-PD.

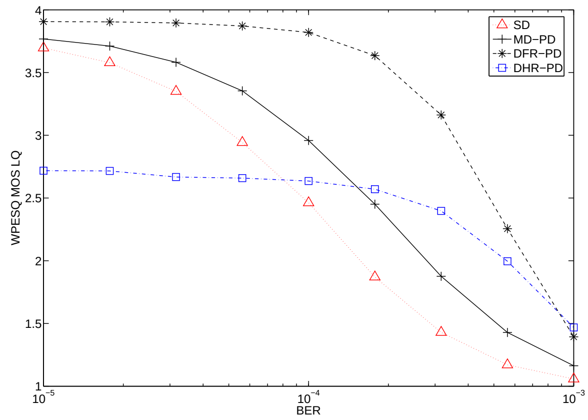

Figure 5: Average WPESQ-MOS values, for changing BER, with compressed packet headers

3.3 Burst Packet Losses

Path diversity is particularly useful in reducing the burstiness of the packet losses. We assume that burst losses are independent of packet sizes. We assume that burst losses are independent of packet sizes because they are usually caused due to phenomena like fading or shadowing in the network or other factors like a link failure. For low rate speech codecs the time required for transmission of a single packet is very small and each packet is transmitted at regular intervals of 10 or 20 milliseconds. The difference between the time required for transmitting say a half-rate codec packet and a full rate codec packet is less than 1 millisecond at a transmission rate of 2 Mbps. So if there is link failure for t ms then the number of packets dropped in this time is same for SD packets and HR packets except in very rare cases where the link failure starts within the time interval of 1 millisecond when a half rate codec would have just finished its transmission but the SD codec needs one millisecond more to finish its transmission. We assume such cases are negligible and the packet loss rate is same for all the codecs we consider.

We model burst losses using a Gilbert model where the channel is modeled using a two-state Markov chain. The channel exists in either a good state or a bad state. No packets are dropped in a good state and all the packets are dropped when the channel is in a bad state. The same tracefiles were used for MD-PD, DHR-PD and DFR-PD. Figure 6 shows the performance of each of the methods for different percentages of average packet losses and an average burst length of 4 (80 ms) packets. We see that the MD-PD method does better than SD but at packet loss rates above 5% the MD-PD performs worse than DHR-PD. This is because a single description of the MD codec is worse than the half-rate codec. When there are burst errors , only one description is received for consecutive packets, the quality of which is below the quality of the HR codec. We can see that DFR-PD performs much better than any other method with only a small increase (around 16% when packet headers are included) in the rate compared to either MD-PD or DHR-PD.

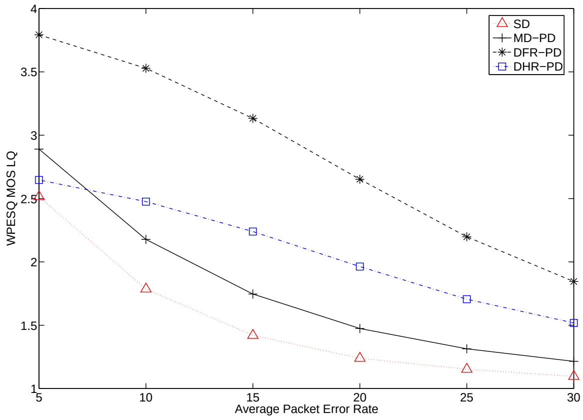

Figure 6: Average WPESQ-MOS values, for burst Errors. Average burst length = 4

4. DISCUSSION

Initially, MD coding seems like a promising method for reliable voice communication over MANETs, but, when packet headers are taken into consideration, we see that a simple path diversity method is not much more inefficient compared to an MD method that uses path diversity. Also, the simple path diversity methods provide a more consistent quality for increasing BERs in the channel.

Another significant point is that the capacity of the network is significantly reduced when any path diversity approach is used instead of sending a single description. Table 4 shows the ratio of the number of bits transmitted for each method to the number of bits transmitted for SD. Observe that without headers MD-PD and DHR-PD have only a small overhead compared to SD whereas DFR-PD pays a huge penalty. When the packet headers are considered, the MD-PD and DHR-PD now have a significant overhead over SD and DFR-PD sends only 16% more bits compared to MD-PD.

Table 4: Ratio of number of bits transmitted for each method

Method	r (without headers)	r (with headers)
SD	1	1
MD-PD	1.125	1.72
DHR-PD	1.0625	1.70
DFR-PD	2	2

5. ACKNOWLEDGMENTS

This work was supported by the California Micro Program, Applied Signal Technology, Dolby Labs, Inc. and Qualcomm, Inc., by NSF Grant Nos. CCF-0429884 and CNS-0435527, and by the UC Discovery Grant Program and Nokia, Inc.

6. REFERENCES

[1] H. Dong, I. Chakeres, H. Dong, E. Belding-Royer, A. Gersho, and J. Gibson, "Selective bit-error checking at the mac layer for voice over mobile ad hoc networks with ieee 802.11," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2004.

[2] A. Servetti and J. D. Martin, "Link-level unequal error detection for speech transmission over 802.11 networks," in *Proc. Special Workshop in Maui (SWIM)*, Jan. 2004.

[3] ——, "Error tolerant mac extension for speech communications over 802.11 wlans," in *IEEE 61st Vehicular Technology Conference, 2005. VTC 2005-Spring.*, 30 May-1 June 2005, pp. 2330 – 2334.

[4] Spectralink. (2006, Mar.) Spectralink voice priority. White Paper. [Online]. Available: http://www.spectralink.com/files/literature /SVP_white_paper.pdf

[5] M. Petracca, A. Servetti, and J. D. Martin, "Voice transmission over 802.11 wireless networks using analysis-by-synthesis packet classification," in *First International Symposium on Control, Communications and Signal Processing*, 2004.

[6] C. Hoene, I. Carreras, and A. Wolisz, "Voice over IP: Improving the quality over wireless LAN by adopting a booster mechanism - an experimental approach," in *SPIE 2001 - Voice Over IP (VoIP) Technology*, 2001, pp. 157–168.

[7] A. Servetti and J. D. Martin, "Adaptive interactive speech transmission over 802.11 wireless LANs," in *Proc. IEEE Int. Workshop on DSP in mobile and Vehicular Systems*, Apr. 2003.

[8] B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, and K. Jarvinen, "The adaptive multirate wideband speech codec (AMR-WB)," in *IEEE Trans. on Speech and Audio Processing*, Nov. 2002.

[9] H. Dong, A. Gersho, V. Cuperman, and J. Gibson, "A multiple description speech coder based on amr-wb for mobile ad hoc networks," in *ICASSP'04*, May 2004.

[10] *Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, ITU-T P.862 Recommendation.

[11] *Proposed modification to draft P.862 to allow PESQ to be used for quality assessment of wideband speech (BT, United Kingdom, and KPN, The Netherlands)*, ITU-T SG12 COM-D007-E Delayed Contribution, Feb. 2001.

[12] V. Barriac, J.-Y. L. Saout, and C. Lockwood. Discussion on unified objective methodologies for the comparison of voice quality of narrowband and wideband scenarios. [Online]. Available: http://portal.etsi.org/stq/presentations2004/090-Mainz_FT_objective_methodologies.doc