

# INTERACTION BETWEEN MULTIPLE DESCRIPTION CODING AND SENSOR NETWORKS WITH FINITE BUFFERS

*E. Baccaglioni*

Dip. di Elettronica - Politecnico di Torino  
Ist. Sup. M. Boella - Multimedia Networking Lab  
10129 Torino, Italy  
enrico.baccaglioni@polito.it

*G. Barrenetxea, B. Beferull-Lozano*

Lab. de Communications Audiovisuelles  
Ecole Polytechnique Federale de Lausanne  
CH-1015 Lausanne, Switzerland  
{guillermo.barrenetxea, baltasar.beferull}@epfl.ch

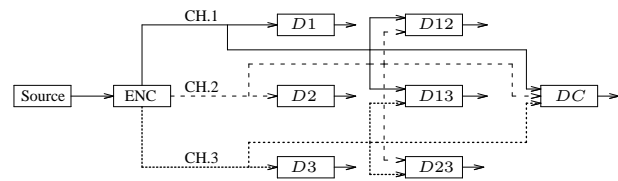
## ABSTRACT

Sensor networks are usually dense networks where the network diversity can be exploited in order to overcome failures. In this paper, we study the use of multiple description techniques in the context of sensor networks where the cause of failures is due to the usual practical constraint of having finite buffers in the sensors, instead of the more traditional case of link failures considered in previous research. Although from a theoretical point of view we observe that the use of more descriptions provides usually better performance, we show experimentally that this is not the case in practice, when real constraints are introduced, such as finite buffers and the presence of header information, necessary for any real application. Our main result is that the optimal number of descriptions, in terms of average distortion, decreases as the fraction of header information increases for a given buffer size.

## 1. INTRODUCTION

Sensor networks are usually composed of a large number of individually-unreliable nodes with many connecting paths between them. Transmissions over these networks are subject to failures such as buffer overflow and node breakdowns resulting in packet losses. This scenario motivates the use of specific coding techniques capable of exploiting network diversity in order to face node unreliability. In this paper, we focus our attention on losses caused by the finite buffer capacity of nodes. Common devices used in sensor networks presents a limited and generally small amount of memory [1] that translates into a limited capacity for the temporary storage of packets. This causes dropping when the rate of packets injected is sufficiently high. In some scenarios, retransmission is not possible due to time constraints or expensive feedback and so coding schemes that make all of the received packets useful can be of great benefit. Multiple Description (MD) coding applies precisely to this situation. As opposed to Single Description (SD) coding, an MD source encoder partitions information into *descriptions* and then sent them over the available paths to the receiver. Depending on the subset of packets that is received, a side decoder computes an estimate of the original source. The quality of the estimate depends on the number of descriptions received but, in contrast to the single-description case, the loss of packets does not lead to a failure. A generic MD coding schema is shown in Fig. 1 for the case of three descriptions.

Several MD coding techniques has been proposed, based on Unequal Error Protection (UEP) [2, 3] and on scalar quantizers



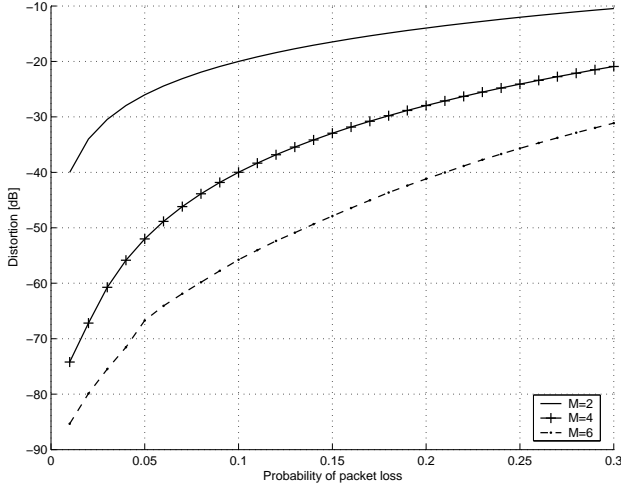
**Fig. 1.** A MD system with three descriptions. Central decoder  $DC$  receives all the descriptions and can reconstruct the finest representation of the source. The other decoders receive only a subset of the transmitted data.

(MDSQ) [4]. However, in all this previous work, only the two-description case is addressed and the interaction between the network conditions and the different practical aspects of building MD codes has not been studied. In this paper, we start by studying the dependence between the number of descriptions and the end-to-end distortion for a point-to-point link and a given fixed probability of failure, using both UEP and MDSQ based techniques. Then, we analyze the application of these coding techniques in a sensor network performing a data-gathering task (simultaneous transmissions case) with finite buffers. We analyze the optimal transmission strategy as a function of practical parameters such as packet sizes and transmission rates in addition to the number of descriptions. Although from a theoretical point of view, the use of more descriptions usually provides a better performance, this is not in general the case when we introduce practical constraints (finite buffers and packets containing both header information and payload). Since the presence of the header information (source, destination, sequence numbers, etc. . .) in each packet containing a description increases the traffic generated for the network, it is shown experimentally that there exists an optimal number of descriptions that depends on the fraction of header information for each value of probability of packet loss.

The rest of the paper is organized as follows: in Section 2, after introducing the model and assumptions, we introduce UEP and MDSQ in a point-to-point channel and in Section 3 we study MD coding in a data-gathering sensor network.

## 2. MDC IN A POINT-TO-POINT CHANNEL

For the analytical study of a MD coding system, we first consider an information source that emits an iid Gaussian signal with zero



**Fig. 2.** Experimental results for a MD-UEP system with a source bitstream of  $L = 48$  bits and no header information.

mean and variance  $\sigma^2$  through a point-to-point channel to the destination node. The point-to-point channel is characterized by a probability  $p$  of losing a packet as it moves from the source to the destination. This parameter depends both on buffer overload in relay nodes and on timeout timers at receivers as we are dealing with a real-time environment. Furthermore, we assume that if the packet reaches the destination, it does not contain bit errors, and, thus, it can be correctly decoded and we also assume that  $p$  does not depend on the length of the packet. We generalize the well-known schemes to deal with an arbitrary number of descriptions both for UEP and MDSQ, denoting with  $M$  their number and with  $R$  their rate.

### 2.1. MDC based on Unequal Error Protection

The information generated by the source is quantized by the MD encoder that generates a progressive bitstream of length  $L = M \cdot R$  bits and then marks it at  $M$  different positions, each one corresponding to the attainment of a distortion level  $D_{s_k}$  ( $1 \leq k \leq M$ ).

Denoting with  $k$  the number of descriptions out of  $M$  correctly received, we evaluate performance using the mean end-to-end distortion:

$$D = \sum_{k=1}^M \binom{M}{k} \cdot D_{s_k} \cdot (1-p)^k \cdot p^{M-k} + \sigma^2 \cdot p^M \quad (1)$$

Our goal is to find the  $M$  values of  $D_{s_k}$  that minimizes, for a given  $p$ , eqn.1.

The technique we used to generate an arbitrary number of descriptions is based on [2] and it optimally tunes the redundancy in each description with linear complexity using a Lagrange multiplier. Its output is then rounded to the nearest integer and the theoretical rate allocation is then guaranteed over a long sequence of samples.

For a source bitstream of  $L = 48$  bits, results are shown in Fig. 2. The end-to-end distortion is plotted as function of the probability of packet loss  $p$ . In this range of values of  $p$ , a 6-description scheme outperforms systems that rely on a smaller

number of packets. The gain is very significant and, for example, for  $p = 0.05$ , the 6-description scheme gains approximately 15 dB over the configuration with four descriptions.

### 2.2. MDC based on index assignment

Together with UEP, another well-known MD coding technique is called Multiple Description Scalar Quantization (MDSQ) and it consists in creating multiple descriptions using  $M$  different quantizers that sample the source of information. In [4] only the two-description case is addressed. Here, we deal with its generalization. A  $M$ -description MDSQ system can be seen as a central quantizer and an *index assignment* in  $M$ -dimension. In fact, the encoder produces, for each source scalar input,  $M$  quantization indices, each one sent over a channel. These indices can be seen as indices of row and column of a  $M$ -dimensional *index assignment hyper-cube* in which each dimension have size  $N = 2^R$  cells. Our goal is to fill this *hyper-cube* with no more than  $2^{MR}$  numbers according to the parameter  $p$ , minimizing the difference in all the *hyper-planes*. Doing so, if only a subset of indices are received, the decoder can reconstruct the best coarse version of the original signal and quality will increase with the number of received descriptions.

We apply the method proposed in [5] to completely fill the *hyper-cube*. As explained in [5], the problem of completely filling it out to minimize the difference in each *hyper-plane* is NP-complete and arrangements of less than  $2^{MR}$  numbers are not addressed. In practice, since we need to fill out only partially the *hyper-cube* in order to achieve trade-offs between central and side distortions, and there does not exist an efficient optimization algorithm, the only possibility is to try all possible ways to fill the *hyper-cube* with less than  $2^{MR}$  numbers, which has a very high computational complexity.

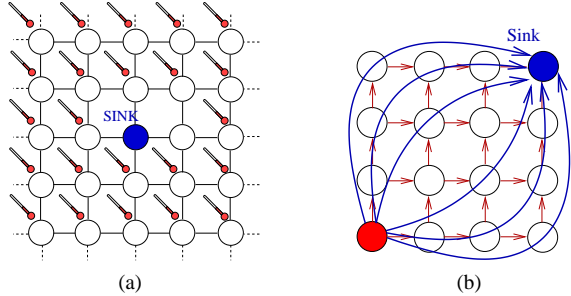
The complexity of dynamically adapt the coder to different operating points discourages a real-time implementation of the index assignment method of [4] in  $M$  dimensions. Anyway, it's worth noticing that, from an analytical point of view, MDSQ, at least for two descriptions, guarantees better performance in terms of mean end-to-end distortion, as proved in [6].

## 3. MDC IN A SENSOR NETWORK

The network model we assumed in the previous section was characterized by the constant probability  $p$  of losing a packet, which is independent of the number of descriptions. Under these assumptions, we showed that the more descriptions we generate, the lower distortion we achieve. However, in many real applications, these assumptions are not valid and a higher number of descriptions does not necessarily lead to a lower distortion.

Suppose we want to measure a Gaussian random field  $X$  on a square area. We assume  $X$  to be uncorrelated in space and time. We uniformly place  $n$  sensor devices that sample the field and send all the information to a single device (sink), which gathers all the information generated by the network (Fig. 3(a)).

We assume that sensors generate samples of  $B$  bits following a Bernoulli distribution with parameter  $\theta$ . Note that  $\theta$  represents also the average information rate generated per device in samples per unit time. To increase the communication efficiency, sensors group  $S$  samples in one single data packet, and send it to the sink. We assume that  $S$  is determined by the application requirements. In practice, in addition to the samples (payload of the packet), data



**Fig. 3.** Network Model: a) Uniform placement of devices that measure a random field and transmit the data to the sink located in the center of the network. b) Routing algorithm: packets are routed using all available paths between the source and the sink with equal probability.

packets contain also a header that includes information such as the device location, sampling times or sequence numbers. We denote by  $H$  the size of the header in bits.

The sensors code the samples and produce  $M$  descriptions per sample. We assume that the total number of bits per sample is fixed. That is, we generate  $M$  descriptions of  $B/M$  bits each so that the information rate transmitted by the network is constant regardless of the number of descriptions.

As  $S$  descriptions are included in the same packet, the total packet size  $K$  in bits is given by  $K = H + S \frac{B}{M}$ .

In terms of communication capability, we assume that the devices can only communicate to their four adjacent neighbors (see Fig. 3(a)). Therefore, to send data to the sink, devices act also as relays for other communications. Packets are routed to the sink using a random routing algorithm that uses all available paths between the source and the sink with equal probability [7]. Note that this routing algorithm allows to exploit the network diversity present in the network, making the use of multiple description coding very convenient (see Fig. 3(b)).

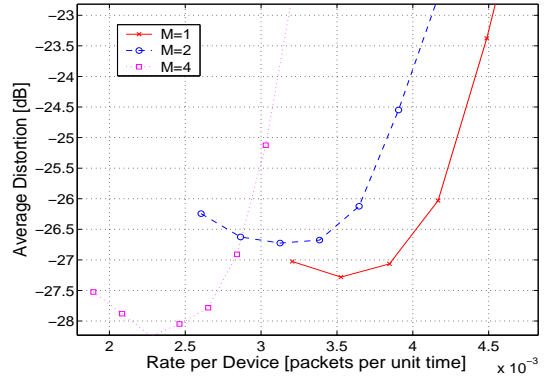
Devices found in real sensor networks generally have a small buffer (queue) for the temporary storage of the packets. We assume that if a node receives a packet and its buffer is full, this packet is lost and no retransmission is attempted. We also assume that transmission time is proportional to the packet size and model the system as a queueing network where each sensor is represented by a queue with a given arrival rate and average service time proportional to the packet size. We further assume that packet losses occurs only due to buffer overflow at the devices. Therefore, the probability  $p$  of losing a packet for a given source is not constant but depends on the network parameters such as the transmission rate  $\theta$ , the packet size  $K$ , and the number of descriptions  $M$ .

With all the information generated by the nodes, we reconstruct the random field  $X$  at the sink. We measure the distortion  $\tilde{D}$  between  $X$  and the reconstructed field  $\hat{X}$  as the average mean-square-error (MSE) per device and unit time. That is:

$$\tilde{D} = \frac{\sum_{i=1}^n (X(i) - \hat{X}(i))^2}{n\theta}. \quad (2)$$

We investigate the optimal transmission strategy that achieves the lowest  $\tilde{D}$ , that is,  $\min_{\{\theta, M, K\}} \tilde{D}$ .

First, note that for fixed values of  $M$  and  $K$ , there exists an optimal transmission rate  $\theta^{\text{opt}}(M)$  that minimizes  $\tilde{D}$ . As we in-



**Fig. 4.**  $\tilde{D}$  as a function of the transmission rate  $\theta$  with  $B = 8$  bits for 1, 2, and 4 descriptions.

crease  $\theta$ , the devices generate more information packets and  $\tilde{D}$  decreases. However, if we continue to increase  $\theta$ , the network becomes congested and devices start losing packets. Fig. 4 shows  $\tilde{D}$  as a function of  $\theta$  in a  $25 \times 25$  network with  $B = 8$  bits per sample,  $H = 36$  bits, and  $S = 20$ .

We examine how  $\tilde{D}$  is affected by the number of descriptions  $M$ . First, note that because of the practical constraints in a real application, every information packet needs to include a header, and therefore, the traffic transmitted through the network increases with the number of descriptions. Moreover, the characteristics of the input traffic of the queues also changes. In MD coding, devices generate several packets at the same time and the traffic becomes more bursty. These two factors contribute to increase the overflow probability and consequently, to increase  $\tilde{D}$ .

On the other hand, when we generate several descriptions, the size  $K$  of each packet is reduced. This packet size reduction decreases network congestion and, equivalently, decreases  $\tilde{D}$ . This can be illustrated with the following example.

Consider just a single  $M/M/1/Q$  queue with a buffer size of  $Q$  packets, average arrival rate  $\lambda$ , and average service time  $\mu$ . Applying  $M/M/1/Q$  formulas, the utilization factor is given by  $\rho = \lambda/\mu$ , and the probability of packet overflow  $p_o$  by:

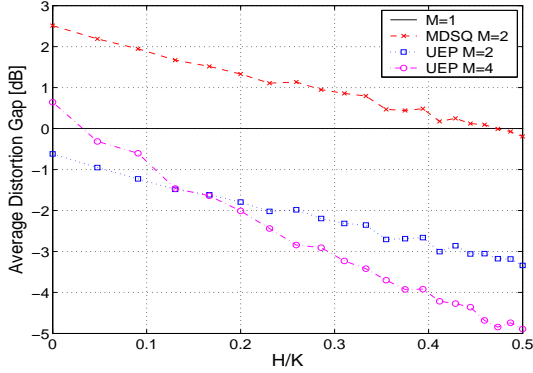
$$p_o = \frac{1 - \rho}{1 - \rho^{Q+1}} \rho^Q. \quad (3)$$

Suppose now that we generate  $M$  times more packets of  $1/M$  the size. Given that the service time is proportional to the packet size, this is equivalent to consider  $\lambda' = M\lambda$  and  $\mu' = M\mu$  and  $Q' = MQ$ . Obviously, the utilization factor does not change, that is,  $\rho' = \rho$ . However, the new probability of packet overflow  $p'_o$  is clearly reduced:

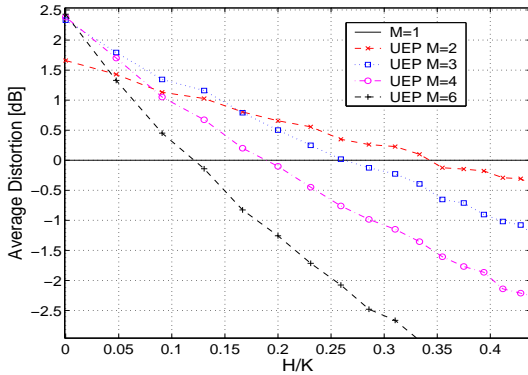
$$p'_o = \rho^{(M-1)Q} \frac{1 - \rho^{Q+1}}{1 - \rho^{MQ+1}} p_o < p_o. \quad (4)$$

Note also that as the average waiting time in the queue is proportional to  $1/\mu$ , packet delay is also reduced. Consequently, to reduce the probability of losing packet due to buffer overflow, smaller packets are preferable.

This suggests that there exists an optimal number of descriptions that depends on the packet size  $K$ , more particularly on the ratio  $H/K$ .



**Fig. 5.** Simulated  $\tilde{D}$  achieved with MD and SD coding for different header-payload ratios and 8 bits per sample.



**Fig. 6.** Simulated  $\tilde{D}$  achieved with MD and SD coding for different header-payload ratios and 12 bits per sample.

We compare now the  $\tilde{D}$  achieved by transmitting at  $\theta^{\text{opt}}(M)$  for different values of  $M$  as a function of the ratio  $H/K$ . Fig. 5 shows the gain in  $\tilde{D}$  achieved by using  $M$  descriptions with respect to SD coding as a function of  $H/K$  in a  $25 \times 25$  network with  $S = 20$  and  $B = 8$  bits per sample for  $M = 1, 2, 4$ . Similarly, Fig. 6 shows the same gain for  $B = 12$  bits and  $M = 1, 2, 3, 4$  and 6.

In the case of  $B = 8$  bits, the optimal strategy for  $H/K < 0.5$  consists in generating 2 descriptions using MDSQ. However, the benefit of using 2 descriptions decreases as the rate  $H/K$  increases due to the extra traffic generated, and when  $H/K > 0.5$ , the lowest  $\tilde{D}$  is achieved by SD coding. Note that due to its exponential complexity, we considered MDSQ only for 2 descriptions.

For  $B = 12$  bits, we considered only UEP coding to generate several descriptions. As we discussed in the previous section, the complexity of MDSQ is exponential not only with  $M$ , but also with  $B$ . Even for the simplest case of  $M = 2$ , each device would have to generate a  $2^{B/2} \times 2^{B/2}$  matrix where  $2^{B/2+1} - 1$  diagonals can be filled. For small  $H/K$  values, the lowest  $\tilde{D}$  is achieved by generating three descriptions, for which a gain of almost 2.5 dB is achieved with respect to SD coding. To generate more than three descriptions does not decrease the distortion: first, we have the practical constraints of creating descriptions with very few bits, and second, there is the penalty due to the extra traffic generated. As the header size increases, the performance of MD degrades due

to this extra traffic. Note that the more descriptions we generate, the more rapidly it degrades. Therefore, the optimal number of descriptions decreases when the ratio  $H/K$  increases. For instance, when  $B = 12$ , we have that for values of  $H/K$  between 0 and 0.15, the optimal strategy consists in generating 3 descriptions. When  $H/K$  is between 0.15 and 0.35, the optimal is  $M = 2$ . Finally, when  $H/K > 0.35$  we use SD coding.

#### 4. CONCLUSIONS

MD coding is a powerful approach to combat packet losses in networks where retransmission is not always possible and it's, thus, suitable for real-time applications. MD coding can be applied over large diversity networks where packets can flow from the source to the destination through many different paths and the probability of losing packets over these paths is not negligible. In this paper, we have studied from an analytical point of view the generalization of two well-known MD techniques to handle many descriptions and we have seen that there exists an optimal number of descriptions to use according to network conditions. Then, we have analyzed the effect of the header size of each description and we have verified that also the header-payload ratio should be carefully taken into account when setting up a Multiple Description Coding system.

#### 5. REFERENCES

- [1] Crossbow Products: Wireless sensor networks, "www.xbow.com/products/new\_product\_overview.htm," .
- [2] R. Puri, T. Kim, and K. Ramchandran, "Multiple description source coding using forward error correction," in *33rd Asilomar Conference on Signals, Systems and Computer*, Oct. 1999, vol. 1, pp. 342–346.
- [3] R. Puri, S. S. Pradhan, and K. Ramchandran, "n-channel symmetric multiple descriptions: New rate regions," in *Proc. Int. Symposium Inform. Th.*, 2002.
- [4] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inform. Th.*, vol. 39, no. 3, pp. 821–834, May 1993.
- [5] Tanya Y. Berger-Wolf and Mitchell A. Harris, "Sharp bounds for bandwidth of clique products," *CoRR*, vol. cs.DM/0305051, 2003.
- [6] V.K.Goyal, "Single and multiple description transform coding with bases and frames," Draft.
- [7] S. D. Servetto and G. Barrenechea, "Constrained random walks on random graphs: Routing algorithms for large scale wireless sensor networks," in *Proc 1st ACM Int. Workshop on Wireless Sensor Networks and Applications (WSNA)*, Sept. 2002.