Reliability Constrained Packet-sizing for Linear Multi-hop Wireless Networks

Ning Wen, and Randall A. Berry
Department of Electrical Engineering and Computer Science
Northwestern University, Evanston, Illinois 60208
Email: {nwen,rberry}@ece.northwestern.edu

***Abstract*— We consider optimizing the packet-sizes and the re-use factor to minimize the delay required to send a message between two nodes in a linear multi-hop wireless network subject to a reliability constraint. Initially assuming no re-use, we give a bound on the required delay. Next, in an infinite system with re-use, we analyze the rate of growth of the delay as a function of the message size. Two cases are considered: one in which packets are decoded/re-encoded on each hop and one in which this is concatenated with an end-to-end outer code. The later is shown to result in lower delays.**

I. Introduction

Consider a wire-line network in which a message is to be divided into several packets of equal size and sent over a sequence of error-free links with transmission rate R. Furthermore, suppose that an entire packet must be received over one link before it may be sent over the next, and that the network is lightly loaded so that whenever this condition is true the packet may be sent. Ignoring any overhead per packet, it is then well-known that the end-to-end delay of the message is minimized by making the packet-size as small as possible so as to benefit from *pipelining*. If overhead is not ignored, there is an optimal packet-size which balances this pipelining effect with the amortization of overhead given by using larger packets.

In this paper, we consider a similar question in the context of a linear multi-hop wireless network. Here, several new issues arise. First, we assume that all nodes transmit in a common frequency-band and so multiple "links" interfere with each other. Also, nodes are precluded from sending and receiving at the same time (i.e., they must satisfy a half-duplex constraint). Finally, we assume that links are not error free and that a node has a constraint on the reliability (i.e., probability of decoding error) at which a message must be obtained. Given this model we consider optimizing the packet-size as well as the schedule of transmissions with the objective of minimizing the total delay for sending a packet from a given source node to a given destination. In this setting for a given transmission schedule packet sizes must now balance the pipelining gains with both the amortization of overhead as well as meeting the reliability constraint.

We formulate a simple linear model for a multi-hop network in which all transmissions are sent hop-by-hop and interfering transmissions are treated as noise. To model the reliability constraint we use an error exponent model derived from the random coding bound for a Gaussian channel. We initially consider a system without spatial re-use, i.e. each link is scheduled in its own time-slot, and derive an upper bound on the end-to-end delay. We then turn to a model with spatial re-use and consider the packet size and re-use distance that either maximize the throughput or minimize the total delay. In the later case, to gain insight we focus on the asymptotic growth of the total delay as the message size L increases to infinity. In this regime, we characterize the optimal number of packets, m for two different coding schemes. In the first, each hop is coded individually. In the second, a concatenated coding scheme is used to add end-to-end coding. In the first case, the optimal number of packets satisfies $m^2 \log(m) = \Theta(L)$, while in the second it satisfies $m^2 = \Theta(L)$. In other words, with the concatenated coding scheme we use smaller packets. The delay under both schemes grows at the same first order rate, but the second order growth is smaller with the concatenated scheme. Finally we conclude with some numerical examples.

Related work includes [1], [2], and [3], which all address linear multi-hop networks. [1] and [2] also consider reliability bounds, but assume interference-free link. In [1], each link is used sequentially, while in [2], nodes are full-duplex. [3] does model interference between nodes and assumes half-duplex transmission as we do here, but focuses on throughput as opposed to delay and reliability.

II. Model

We consider a one-dimensional model, where all nodes are regularly placed on a one-dimensional line and labeled by an integer. One node x is assumed to have L nats[1] of data to send to another node y. Let D be the distance between node x and y, and let $H-1$ be the number of nodes between x and y, each of which is assumed to be a relay for the message (i.e. the number of "hops" is H). To simplify the analysis, we assume the queuing delay on each hop is zero. This is reasonable assuming that the given flow has higher priority over other flows in the network.

All nodes are assumed to transmit in the same frequency band (with normalized bandwidth of 1) and treat any interfering transmission as noise. The channel between any pair

This research was supported in part by DARPA under grant W911NF-07-1-0028, and by NSF under grant CNS-0626558 and by ARO under grant W911NF-06-1-0339.

[1] In order to simplify notation, we use *nat* as the unit for information in this paper.

of nodes is modeled by a distance dependent path-loss plus additive Gaussian noise. Furthermore, we assume that the nodes employ a regular TDM-schedule of length K so that in time-slot t, the nodes $nK + (t \mod K)$ are allowed to transmit, for $n = \dots, -1, 0, 1, \dots$. Figure 1 illustrates this space-time reuse model.

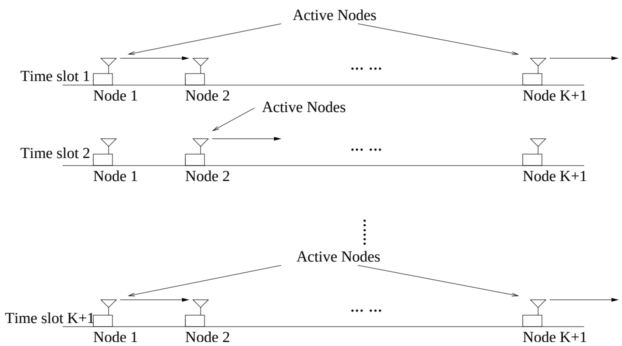

Fig. 1. Space-time reuse scheme.

To begin consider a simple model where all the nodes have the same transmission rate. Assume the L nats of information are transmitted in m equal-size packets, and there are h additional nats of overhead in each packet. Let $R(H, K)$ denote the transmission rate in terms of *nats per channel use* under a given H and K and assume that all transmissions are reliable. It follows that $\frac{\frac{L}{m}+h}{R(H,K)}$ is the length of the time-slot for one packet. The end-to-end delay in channel uses is then

$$D(H, K) = Km\frac{\frac{L}{m} + h}{R(H, K)} + (H - K)\frac{\frac{L}{m} + h}{R(H, K)}, \tag{1}$$

where the first term is the time for the source to send all the L nats over the first K hops, and the second term is the time required for the last packet to traverse the remaining $H - K$ hops. Essentially, this is a pipelining calculation as described in the introduction, only now a packet must traverse K hops before a new packet may be transmitted.

If H and K are fixed, and we ignore the integer constraint on m, then the delay is minimized if m takes the value:

$$m^*(H, K) = \sqrt{\frac{(H - K)L}{Kh}}, \tag{2}$$

and the optimal delay is:

$$D^*(H, K) = \frac{\left(\sqrt{KL} + \sqrt{(H - K)h}\right)^2}{R(H, K)}. \tag{3}$$

Notice that if the overhead h is negligible, then the optimal m goes to infinity, and so the optimal packet size of $\frac{L/m}{R(H,K)}$ channel uses goes to zero. We are interested in the case where each packet must also meet a reliability constraint. This, in addition to overhead considerations, will prevent the use of arbitrarily small packets. Specifically we assume that the end-to-end message error probability for delivering the L nats must be no greater than a constant η. Given this constraint, we then want to minimize the end-to-end delay.

Assuming that each node uses a random Gaussian code, then from [4], [5], the one hop block error probability P_{hb} of a code with a block-length of N_{hb} channel uses is bounded by:

$$P_{hb} \le \exp(\rho L_b - N_{hb}(E_0(\rho, SINR))), \tag{4}$$

for any $\rho \in [0, 1]$, where L_b is the number of information nats contained in one block, $SINR$ is the Signal-to-Noise-plus-Interference ratio at the receiver, and $E_0(\rho, SINR)$ is the error exponent determined by ρ and $SINR$. For a complex Gaussian channel with unit bandwidth, a simple expression for the error exponent is given by:

$$E_0(\rho, SINR) = \rho \log\left(1 + \frac{SINR}{1 + \rho}\right). \tag{5}$$

Given an upper bound η_{hb} on the block error probability P_{hb} for a single hop, the minimum N_{hb} satisfying

$$N_{hb}\rho \log\left(1 + \frac{SINR}{1 + \rho}\right) \ge \rho L_b - \log \eta_{hb} \tag{6}$$

is the minimum delay for sending a block of L_b nats over that hop for which we can use (5) to guarantee that the reliability constraint is met. In the rest of the paper, we use this value of N_{hb} as the minimum delay for each hop.

III. Delay without Spatial Reuse in Finite Length System

Now we return to problem of sending L nats of data between two nodes x and y over H equal length hops, but include the reliability constraint. We assume that the distance D between the source and destination is normalized[2] to 1. Furthermore, there is no other interference present. In this section, we consider the simplest case where only one transmission is possible in each time slot. Let N_0 denote the noise power spectrum density. We assume the transmission power is P for each transmitter.

We consider the optimal number of hops H and the number of blocks m containing the original L nats. Before going into the details, we introduce the following basic result.

Lemma 1: In an optimized system, the delay for each hop is the same for each packet.

This follows directly from the assumptions that each packet has the same size and that the channel for each hop is the same and thus so is its SINR.

Thus we assume that the delay is the same for each hop and we denote this by N_{hb}. Let $SINR = \frac{P}{N_0}H^\alpha$ be the $SINR$ for a hop, where α is the path-loss factor. Then, (6) can be rewritten as:

$$N_{hb}\rho \log\left(1 + \frac{\frac{P}{N_0}H^\alpha}{1 + \rho}\right) \ge \rho\left(\frac{L}{m} + h\right) - \log \eta_{hb}, \tag{7}$$

where η_{hb} is the per hop, per block error constraint. Assuming no retransmissions, and that coding is done independently over

[2]Changing this distance is equivalent to changing the transmission power. Since our results are suitable for any power, we use the normalized distance to simplify our notation.

each hop then η_{hb} and the end-to-end error constraint η will satisfy

$$1-\eta=(1-\eta_{hb})^{mH}. \quad (8)$$

Considering that $\eta<1$ and $mN\geq 1$, we have

$$(1-\frac{\eta}{mH})^{mH}\geq 1-\eta. \quad (9)$$

This implies that setting $\eta_{hb}=\frac{\eta}{mH}$ guarantees the end-to-end message error probability is less than η.

Let $N=N_{hb}\times H\times m$ denote the total delay, then the minimum N satisfying

$$N\geq \frac{HL+Hmh-\frac{Hm}{\rho}\log\frac{\eta}{Hm}}{\log\left(1+\frac{\frac{P}{N_0}H^{\alpha}}{1+\rho}\right)} \quad (10)$$

can be guaranteed to satisfy the reliability constraint.

Proposition 1: Without an integer constraint on H, the minimum total delay N^* is upper bounded by

$$N^*\leq \frac{H^*L+H^*h-\frac{H^*}{\rho}\log\frac{\eta}{H^*}}{\log\left(1+z(\alpha)\right)}, \quad (11)$$

where $H^*=(\frac{(1+\rho)z(\alpha)}{P/N_0})^{\frac{1}{\alpha}}$, and $z(\alpha)$ is the solution to $(1+z^{-1})\log(1+z)=\alpha$.

Proof: Notice that H and m are positive integers. The right-hand-side of (10) is monotonically increasing with m. Thus, for any given H, the m that minimizes (10) is 1. Note that

$$\frac{HL+Hh-\frac{H}{\rho}\log\eta}{\log\left(1+\frac{\frac{P}{N_0}H^{\alpha}}{1+\rho}\right)}<\frac{HL+Hh-\frac{H}{\rho}\log\frac{\eta}{H}}{\log\left(1+\frac{\frac{P}{N_0}H^{\alpha}}{1+\rho}\right)}. \quad (12)$$

Minimizing, the left-hand side of (12) with respect to H yields the desired result. ■

Note that H^* is independent of L, h, η, and H^* decreases as $\frac{P}{N_0}$ grows.

IV. Infinite Length System

Now we consider the situation where simultaneous transmissions are enabled so that the $SINR$ includes the effect of interference. To calculate the interference, knowledge of the sets of transmitters in each time-slot is necessary, which complicates the problem. To simplify this we assume there are infinitely many nodes regularly placed in a one-dimensional line and that each node in the line always has traffic to send. Thus, for a given transmission schedule each node will see the same SINR.

In this section, we normalize the distance between *adjacent* nodes to be[3] 1 and assume the number of hops H between the source node x and the destination node y is a constant.

The nodes follow a space-time reuse schedule as in Fig. 1, so that for a schedule of length K the distance between two adjacent transmitters is K hops. Each user has an average power constraint of[4] P and so transmits with power KP when scheduled. We are still assuming that the message from x can be sent to y via multi-hop transmission without queuing delay and are primarily interested in the case where the message size L is large.

A. Throughput Optimization

Before considering the optimal delay problem, we address the related problem of optimizing the end-to-end throughput given a fixed block size of L_b nats and an end-to-end error constraint of η. Let N_{hb} be the one-hop minimum delay for L_b nats of information, which meets the corresponding per-hop reliability constraint in (6) The average throughput can then be written as

$$R=\frac{L_b}{KN_{hb}}. \quad (13)$$

We next consider maximizing this over L_b and K.

Let $\gamma(\frac{N_0}{P},K)$ be the received $SINR$ at each node, where

$$\gamma(\frac{N_0}{P},K)=\frac{1}{\frac{N_0}{P}\frac{1}{K}+\sum_{i=1}^{\infty}(iK+1)^{-\alpha}+\sum_{j=1}^{\infty}(jK-1)^{-\alpha}}. \quad (14)$$

Now, the reliability constraint (6) can be written as

$$N_{hb}\log\left(1+\frac{\gamma(\frac{N_0}{P},K)}{1+\rho}\right)\geq L_b-\frac{1}{\rho}\log\frac{\eta}{H}. \quad (15)$$

Substituting (15) into (13) yields

$$R\leq \frac{L_b}{L_b-\frac{1}{\rho}\log\frac{\eta}{H}}\frac{\log\left(1+\frac{\gamma(\frac{N_0}{P},K)}{1+\rho}\right)}{K}. \quad (16)$$

The right-hand side of (16) can be decomposed into two parts, denoted by R_L and R_K, respectively, Assuming a fixed value of ρ, the first part $R_L=\frac{L_b}{L_b-\frac{1}{\rho}\log\frac{\eta}{H}}$ only depends on L_b, and the second part $R_K=\frac{1}{K}\log(1+\frac{\gamma(\frac{N_0}{P},K)}{1+\rho})$ only depends on K.

It is clear that R_L is maximized if L_b takes the largest possible value, and a finite K maximizes R_K. To gain some insight, we consider the optimal K when $\frac{N_0}{P}$ takes extreme values. In the low $SINR$ regime, where $\frac{N_0}{P}$ goes to ∞, $\gamma(\frac{N_0}{P},K)$ goes to 0 and optimal K is the smallest possible K. On the other hand, if $\frac{N_0}{P}$ goes to 0, then the system is in interference limited region and the optimal K is a bounded constant determined by the system parameters.

B. Optimal Delay

The analysis in Section IV-A shows that longer block lengths are preferred to maximize throughput. However, this prohibits pipelining and introduces longer delay. Now we return to the problem of minimizing the delay to send a message of L nats of information over H hops. As before this

[3]Notice that this is a different assumption than the one in the previous section.

[4]Notice this is also a slightly different normalization than in the previous section.

message can be divided m blocks of equal size $L_b = \frac{L}{m} + h$, where h denotes the additional overhead needed per packet. The delay D_1 of sending one block over one hop must now satisfy

$$D_1 \geq \frac{\left(\frac{L}{m} + h\right) - \frac{1}{\rho} \log \frac{\eta}{mH}}{\log\left(1 + \frac{\gamma(\frac{N_0}{P}, K)}{1+\rho}\right)}, \tag{17}$$

to ensure the reliability constraint is met. Note that D_1 is also the length of one time-slot.

From the discussion in Section II, the number of time slots from x sending out the first block until y receives the last block is $H + (m-1)K$. Thus the total delay should satisfy

$$D \geq \frac{(H + (m-1)K)\left[\left(\frac{L}{m} + h\right) - \frac{1}{\rho} \log \frac{\eta}{mH}\right]}{\log\left(1 + \frac{\gamma(\frac{N_0}{P}, K)}{1+\rho}\right)}. \tag{18}$$

To gain insight into the parameters which optimize this, we next consider an asymptotic regime, where L goes to ∞.

V. Asymptotic Analysis

In this section, we study the asymptotic behavior of the optimal parameters m, ρ and K as the total information L goes to ∞.

A. Optimal orders of ρ and m

The right-hand side of (18) can be rewritten as

$$\begin{aligned} D(\rho, m) \quad = \quad & \frac{1}{\log\left(1 + \frac{\gamma(\frac{N_0}{P}, K)}{1+\rho}\right)} \Bigg(KL + b_1 \frac{L}{m} \\ & + K\frac{1}{\rho} m \log m + K(h + b_2)m \\ & + b_1 \frac{1}{\rho} \log m + b_1(h + b_2) \Bigg), \end{aligned} \tag{19}$$

where $b_1 = H - K$, and $b_2 = -\log \frac{\eta}{H}$. We consider the behavior of this as $L \to \infty$ for a fixed K.

Proposition 2: Let ρ^* and m^* minimize (19) over $0 \leq \rho \leq 1$ and $m \geq 1$. If $L \to \infty$, then ρ^* and m^* satisfy $\rho^* \to 0$, $m^* \to \infty$, and $\frac{1}{\rho^*} m^{*2} \log m^* = \Theta(L)$.

Proof: On the right-hand-side of (19), note that $\left(1 + \frac{\gamma(\frac{N_0}{P}, K)}{1+\rho}\right)$ is bounded, no matter what value ρ is used. The highest order term is KL, and all the other terms are of smaller orders than L. This implies that $m \to \infty$, and $\frac{1}{\rho} m \log m = o(L)$. The coefficient of the highest order term, KL, is $\left(1 + \frac{\gamma(\frac{N_0}{P}, K)}{1+\rho}\right)^{-1}$. This is minimized when $\rho \to 0$. Therefore, $\rho \to 0$, $m \to \infty$, and $\frac{1}{\rho} m \log m = o(L)$.

Now the candidates for the second highest order terms are $b_1 \frac{L}{m}$ and $K \frac{1}{\rho} m \log m$. These two terms should be asymptotically equivalent to minimize the total delay. Thus

$$b_1 \frac{L}{m} \asymp K \frac{1}{\rho} m \log m, \tag{20}$$

which yields the desired result. ■

Proposition 2 shows that $\frac{1}{\rho^*} m^{*2} \log m^*$ increases linearly with L, if ρ is fixed. Suppose that there is no reliability constraint, *i.e.* the transmission rate is determined by $SINR$ and is independent of m. In this case, from (2), it follows that m^* should increase at the order of $\Theta(L^{\frac{1}{2}})$. This implies that m grows faster without a reliability constraint than when one is present. In other words, with the reliability constraint, an order larger block-size is preferred.

B. Optimal K

From the proof of Proposition 2, the highest order term of $D(\rho, m)$ as $L \to \infty$ is

$$D^1(\rho, m) = \frac{KL}{\log\left(1 + \frac{\gamma(\frac{N_0}{P}, K)}{1+\rho}\right)}.$$

Furthermore, ρ^* is shown to go to 0 in Proposition 2. Thus the optimal K is given by

$$K^* = \arg\min_k \frac{K}{\log\left(1 + \gamma(\frac{N_0}{P}, K)\right)}. \tag{21}$$

As $K \to \infty$, $\frac{K}{\log\left(1+\gamma(\frac{N_0}{P}, K)\right)} \to \infty$. Thus the optimal K is a bounded constant. Notice that (21) is the inverse of R_K in Section IV-A, and so the same results are valid for the extreme values of K in high and low SINR regimes.

VI. Concatenated Coding Scheme

Previously, we assumed a message containing L nats was divided into m blocks, and the message was successfully received when there is no error for every block on every hop. Next, we relax this assumption and consider a concatenated coding scheme as in [6] and [2], in which an outer code is also used to correct missing packets which do not arrive at the destination.

Assume that the inner code length is N_i channel uses, and the outer code length is N_{out}. Assume the dimensionless rate of the outer code is r, then

$$N_{out} = \frac{m}{r} N_i. \tag{22}$$

The error probability per block per hop, is here given by

$$P_{bh} \leq \exp\left(\rho\left(\frac{L}{m} + h\right) - N_i \rho \log\left(1 + \frac{\gamma(\frac{N_0}{P}, K)}{1+\rho}\right)\right). \tag{23}$$

The end-to-end block error, which is also the inner code error, can then be upper bounded as

$$P_b \leq H P_{bh} \leq \exp\{-N_i E_i\}, \tag{24}$$

where $E_i = \rho \log\left(1 + \frac{\gamma(\frac{N_0}{P}, K)}{1+\rho}\right) - \frac{\rho L}{m N_i} - \frac{\rho h}{N_i} - \frac{\log H}{N_i}$ can be interpreted as the error exponent of the inner code.

Based on [6], the error exponent of the outer code can be expressed as

$$E_o(R_o) = \max_{r R_i = R_o} (1 - r) E_i(R_i), \tag{25}$$

where R_i and R_o are the rate for the inner and outer codes, respectively. And $R_i = \frac{\frac{L}{m}+h}{N_i}$ and $R_o = \frac{L+mh}{N_{out}}$.

Proposition 3: Let ρ^* and m^* minimize the total delay over $0 \le \rho \le 1$ and $m \ge 1$. With the concatenated coding scheme, if $L \to \infty$, then the optimal ρ^* and m^* satisfy $\rho^* \to 0$, $m^* \to \infty$, and $\frac{1}{\rho^*}m^{*2} = \Theta(L)$.

Proof: The total message error can be expressed as[5]

$$\begin{aligned} &P_e \\ \le\ & \exp\left\{-\max_{rR_i=R_o}\left[(1-r)\left(\rho\log\left(1+\frac{\gamma(\frac{N_0}{P},K)}{1+\rho}\right)\right.\right.\right. \\ &\left.\left.\left.-\frac{\rho L}{mN_i}-\frac{\rho h}{N_i}-\frac{\log H}{N_i}\right)\right]N_{out}\right\} \\ =\ & \exp\left\{-\left(\sqrt{N_{out}\rho\log\left(1+\frac{\gamma(\frac{N_0}{P},K)}{1+\rho}\right)}\right.\right. \\ &\left.\left.-\sqrt{\rho L+\rho mh+m\log H}\right)^2\right\}, \end{aligned} \tag{26}$$

where the optimal r in the final step is

$$r^* = \sqrt{\frac{\rho L+\rho mh+m\log H}{N_{out}\rho\log\left(1+\frac{\gamma(\frac{N_0}{P},K)}{1+\rho}\right)}}. \tag{27}$$

To satisfy the end-to-end error requirement η, we have

$$\sqrt{N_{out}\rho\log\left(1+\frac{\gamma(\frac{N_0}{P},K)}{1+\rho}\right)} \ge \sqrt{\rho L+\rho mh+m\log H}+\sqrt{-\log\eta}. \tag{28}$$

Using the same argument as the previous section, the total end-to-end delay is

$$D(\rho,m) = (H+(m-1)K)\frac{rN_{out}}{m}. \tag{29}$$

Substituting (27) and (28) into (29), and using the same technique as in Proposition 2, yields the result. ■

Notice that the total end-to-end delay still grows linearly with L. However, the second highest order terms now grow at a slower rate than the scheme without concatenated coding, which implies that this scheme will have a smaller end-to-end delay. As $L \to \infty$, the difference between the delays of the two schemes also grows to infinity.

VII. Numerical Result

To illustrate the asymptotic analysis, some numerical results are provided in this section. Figure 2 shows the growth order of the minimal delays for both the schemes with and without concatenated coding. It is clear that the highest order term dominates the total delay, when the amount of information goes to infinity. It also illustrates the behavior of the optimal K for different L. Figure 3 shows the optimal ρ and m as a function of L. The results shown in these two figures are consistent with our analysis. In all of above figures, $\alpha = 3$, $\eta = 0.001$, $N_0/P = 1$, $h = 10$, and $H = 10$.

[5]Notice E_o is the actual error exponent when r is close to 1. We use E_o to calculate the error probability since r indeed approaches 1 when L is large.

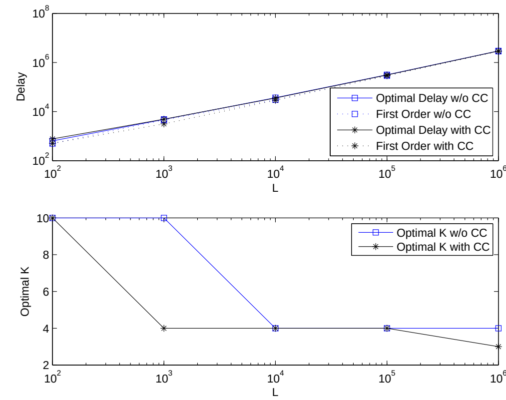

Fig. 2. Minimum delays and optimal K's.

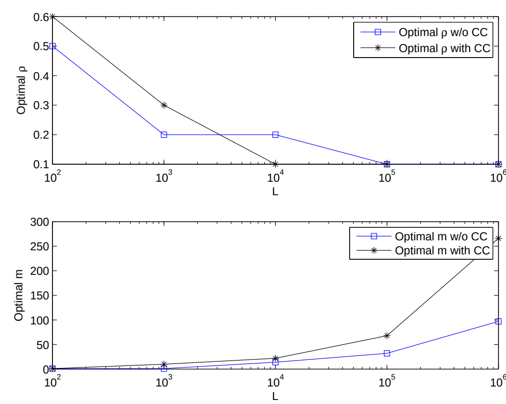

Fig. 3. Optimal ρ's and m's.

References

[1] O. Oyman, "Reliability bounds for delay-constrained multi-hop networks," in *44th Annual Allerton Conference*, Illinois, Sept. 2006.

[2] W. Zhang and U. Mitra, "Multihopping strategies: An error-exponent comparison," in *IEEE International Symposium on Information Theory (ISIT)*, Nice, France, June 2007.

[3] M. Sikora, J. N. Laneman, M. Haenggi, D. J. Costello, and T. Fuja, "On the Optimum Number of Hops in Linear Ad Hoc Networks," in *IEEE Information Theory Workshop (ITW'04)*, San Antonio, TX, Oct 2004.

[4] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: John Wiley & Sons, Inc., 1968.

[5] I. E. Telatar and R. G. Gallager, "Combining queueing theory with information theory for multiaccess," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 6, pp. 963–969, August 1995.

[6] G. D. Forney, *Concatenated Codes*. M.I.T. Press, 1966.