Solutions to final exam Spring 2006

Problem 1 Consider a relay network with node 1 being the source node, node 2 the relay node, and node three the terminal node. Suppose for $i \in \{1, 2\}$, that node i sends a coded symbol in each time slot, and each transmission is received without error at the next node with some fixed probability z_i, and is erased with probability $1 - z_i$. The symbols sent by node 1 are generated by LT coding, based on K input symbols in $\{0, 1\}^L$ for $L >> 1$, and a robust soliton degree distribution. Node 2 in each slot sends the coded symbol that it most recently received from node 1. Node 3 is able to signal back to the other two nodes when it has received $K(1 + \epsilon)$ symbols, allowing it to decode successfully, but no other feedback is available.

(a) What is the probability that node 3 receives at least one copy of a given output symbol? (Hint: This is the throughput, measured in symbols per slot, if ϵ and packet headers are ignored. If $z_1 = z_2 = 1/2$, the throughput is 1/3.)

(b) Suppose the strategy is changed at node 2, so that node 2 continues to send one of the packets it received in each slot, but the symbols is chooses to send can be chosen differently. Give a choice rule leading to a higher information rate than the one given. In particular, show that throughput 3/8 is possible if $z_1 = z_2 = 1/2$ (again, ignoring factors arbitrarily close to one).

(c) What information rate is achievable if node 2 is allowed to combine the symbols it receives to make new ones? Explain.

Solution (a) Method A (Lei Ying) A symbol will be received at node 3 j slots after the slot it is transmitted by node 1 and no sooner, with probability $z_1(1 - z_1)^{j-1}(1 - z_2)^{j-1}z_2$. Summing over $1 \leq j < \infty$ yields that a symbol sent by node one is eventually received with probability $\frac{z_1 z_2}{1-(1-z_1)(1-z_2)}$.

Method B No copies will be received if the symbol does not make it to node 2. If the symbol does make it to node 2, no copies will be received if all transmissions of it by node 2 are erased. Given that node 2 receives a symbol, the number of times node 2 transmits it has the geometric probability distribution $(z_1(1 - z_1)^{j-1} : j \geq 1)$. Thus,

$$\begin{aligned} P[\text{symbol does not reach node 3}] &= (1 - z_1) + z_1 \left[\sum_{j=1}^{\infty} z_1(1 - z_1)^{j-1}(1 - z_2)^j \right] \\ &= 1 - z_1 + \frac{z_1^2(1 - z_2)}{1 - (1 - z_1)(1 - z_2)} \end{aligned}$$

The throughput is $1 - P[\text{symbol does not reach node 3}]$.

(b) The throughput would be larger if in each slot node 2 sent a coded symbol that it previously sent the least number of times. If $z_1 = z_2 = 0.5$, then in the long run, node 2 will be able to send each of the symbols it receives twice, so that each will have a 3/4 chance of being received at node 3. This yields throughput rate 3/8.

(c) Throughput $\max\{z_1, z_2\}$ can be achieved by network coding.

Problem 2 (Variation of Mitzenmacher and Upfal, problem 5.11) A system initially with n balls and n bins operates in rounds. In each round, each remaining ball is thrown into a bin. Any ball landing in a bin which no other ball falls into during the round, is removed. The other balls are carried over to the next round.
(a) Let $g(b)$ denote the number of balls remaining at the end of a round, given there are b balls at the start of a round. Find $g(b)$ and show that $g(b) \leq b^2/n$.
(b) Let $x_0 = n$ and $x_{j+1} = g(x_j)$ for $j \geq 1$. Thus, x_j refers to the number of balls left after j rounds, under the approximation that the outcome of each round exactly follows expectations. Let $\tau = \min\{j : x_j \leq 1\}$. Show that $\tau = O(\ln \ln n)$ as $n \to \infty$.
(c) Suppose the number of balls at the beginning of a round is a random variable B, which is Poisson distributed with mean b. Describe the probability distribution of the number of balls left at the end of the round.
(d) (10 points extra credit) Prove that if c is sufficiently large and σ is the (random) number of rounds required until all balls are removed, then $P\{\sigma \geq c \ln \ln n\} \to 0$ as $n \to \infty$.

Solution (a) Any given slot will have a single packet in it with probability $\frac{b}{n}(1-\frac{1}{n})^{b-1}$, so that the mean number of balls removed is $b(1-\frac{1}{n})^{b-1}$. Thus, the mean number remaining is $g(b) = b - b(1-\frac{1}{n})^{b-1}$. Using the inequality $(1-\alpha)^k \geq 1 - \alpha k$ yields that $g(b) \leq b - b(1 - \frac{b-1}{n}) = \frac{b(b-1}{n} \leq \frac{b^2}{n}$.
(b) Since $\lim_{n\to\infty}(1-\frac{1}{n})^{n-1} = e^{-1}$, if follows that $x_1 = g(n) = n(1-(1-\frac{1}{n})^{n-1}) \leq (0.7)n$ for n sufficiently large. For such n, by part (a), $x_2 \leq (0.7)^2 n$, $x_3 \leq (0.7)^4 n$, and by induction, $x_k \leq (0.7)^{2^{k-1}} n$. Thus, $x_k \leq 1$ if $(0.7)^{2^{k-1}} \leq \frac{1}{n}$ or equivalently $2^{k-1} \geq \frac{\ln n}{-\ln 0.7}$, or equivalently $k \geq \ln(\frac{\ln n}{-\ln 0.7})/\ln 2 = O(\ln \ln n)$.
(c) It X denotes the number of balls falling in a given bin, then X has the Possion distribution is mean $\frac{b}{n}$. The number of balls falling in the bin that remain in the bin for the next round is $X - I_{\{X=1\}}$. Thus, if $Z_1, \ldots, Z_n$ are independent and each has the same distribution as $X - I_{\{X=1\}}$, then $W(b) = Z_1 + \cdots + Z_n$ has the same distribution as the number of balls remaining at the end of the slot.
(d) Here is a sketch. Part (c) can be used to show that for any $\epsilon > 0$, there is a $\delta > 0$ and constant c so that $P\{W(b) \leq (1+\epsilon)g(b)\} \leq \frac{c}{b^\delta}$ for all b. Let X_k denote the number of balls remaining after k rounds. It can then be shown that if $\tau = \min\{k : X_k \leq \Gamma\}$ for a sufficiently large value of Γ, then $P\{X_j \leq (1+\epsilon)g(X_{j-1})$ for $1 \leq j \leq \tau\} \to 1$ as $b \to \infty$. Thus, if (y_j) is the sequence with $y_0 = n$ and $y_{j+1} = (1+\epsilon)g(y_j)$, and if $T = \min\{j : y_j \leq 1\}$, then $P\{X_{T+1} = 0\} \to 1$ as $b \to \infty$. Arguing as in part (b) it can be shown that $T = O(\ln \ln n)$.