

# CMSC 858S: Algorithms in Networking

Fall 2004

Ungraded Homework Assignment #2, handed out Oct. 4, 2004

**Note:** We will have *ungraded* homework assignments such as this one, as well as ones that will be graded. I will post the solutions for all the assignments some time after they are handed out. You will get the most out of this course if you do your best to solve all the homework problems (whether they are graded or not) in collaboration with your group-members. In addition, I hope the collaboration will be a rewarding experience in co-operative work for you; please let me know if you have found simpler or more elegant solutions than the ones handed out.

The suggested deadline by which to finish this assignment is October 16th; since this assignment is ungraded, you don't need to turn it in – just compare your solutions with the solutions I give. This homework mainly deals with techniques to handle *correlations*. All of problems 1-3 may be quite challenging.

1. We encountered the following problem while discussing Probabilistic Resilient Multicast. Suppose a node  $u$  chooses a node  $v$  from among  $n$  nodes at random, and chooses to send a message with probability  $p$  to  $v$ . If *both* of these events happen, we say that “ $u$  communicated with  $v$ ”; note that for any given node  $v$ , the probability that  $u$  communicated with  $v$  is  $q = p/n$ . Let  $S$  be some set of nodes, and let  $w$  be a node not lying in  $S$ . Suppose we are given that  $u$  **did not** communicate with any node in  $S$ . Given this, is the conditional probability of  $u$  communicating with  $w$  greater, equal, or less than  $q$ ? (**Hint:** Induction on the size of  $S$  may help.)

2. We showed in class that under our assumptions of suitably high degree etc., all surviving non-leaves get the data with high probability. We then claimed that at least an  $(1-\epsilon)\cdot(1-\delta)\cdot(1-o(1))$  fraction of the surviving leaves also get the data with high probability; prove this claim.

3. The following correlational inequality [2, 1] involving a “balls and bins” experiment is often useful: *Suppose we throw  $m$  balls independently at random into  $n$  bins, each ball having an arbitrary distribution. Let  $B_i$  be the random variable denoting the number of balls in the  $i$ th bin. Then for any  $t_1, \dots, t_n$ ,  $\Pr(\bigwedge_{i=1}^n B_i \geq t_i) \leq \prod_{i=1}^n \Pr(B_i \geq t_i)$ .*

In the experiment where we throw  $n$  balls uniformly at random and independently into  $n$  bins, let  $L_n$  denote the maximum number of balls in any bin. Use the above correlational inequality to show the following: “there is a constant  $c > 0$  such that with probability tending to 1 as  $n \rightarrow \infty$ ,  $L_n \geq c \log n / \log \log n$ ”. (Thus, combined with our discussion in class, we see that  $L_n = \Theta(\log n / \log \log n)$  with high probability.)

4. Read paper [Dub98] in the class references webpage (the paper by Dubhashi on Martingales). *The suggested deadline to complete this reading is by mid-November. I will assume for the final exam that you have done this reading.*

## References

- [1] D. Dubhashi and D. Ranjan. Balls and bins: A study in negative dependence. *Random Structures & Algorithms*, 13(2):99–124, Sept 1998.
- [2] C. L. Mallows. An inequality involving multinomial probabilities. *Biometrika*, 55:422–424, 1968.