

Problem Set 2, Due Tuesday, February 12, 2008 (in class for non-SCPD students)

Each pair of students only needs to submit one set of answers.

1. Consider a directed 4-cycle, i.e. node A has an edge to node B, B has an edge to C, C has an edge to D, and D has an edge to A.
 - (a) Compute the naive PageRank of node A. [2pts]
 - (b) Now add an edge from node A to node D. Recompute the naive PageRank of node A. Can you give a plain English explanation of this phenomenon? [5pts]
 - (c) Assume that each node is allowed to add additional outgoing edges to the original 4-cycle. More precisely, node A is allowed to have (i) an edge to B alone, or (ii) edges to B and C, or (iii) edges to B and D, or (iv) edges to B, C, and D. The strategy space for B, C, and D are similarly defined. Each node must have an edge pointing to the next node in the cycle, and is allowed to have as many additional outgoing edges as it wants. Edges from a node to itself are not allowed. Also, for any two nodes X and Y, there can be at most one edge from X to Y and at most one edge from Y to X. Each node wants to maximize its own naive PageRank. Prove that the 4-cycle does not represent a Nash equilibrium. [3pts]
 - (d) Suppose each node adds an edge to its predecessor in the 4-cycle, i.e. A adds an edge to D, B adds an edge to A, C adds an edge to B, and D adds an edge to C. Prove that this graph (known as a bidirected 4-cycle) represents a Nash equilibrium in the game outlined in part (c). [5pts]
 - (e) Find another Nash equilibrium in this game. [5pts]
2. Give a brief explanation of why a free equivalent of eBay (such as FreeBay) might not be stable. Comment on how (and to what extent) Craigs' List is able to overcome this problem. [10pts]
3. Consider a slight variant of the model for long tail phenomena described in class: Suppose the t -th new user in a system introduces and consumes one new product. This user also consumes k existing products, each with probability proportional to the number of times a product has been consumed already. Let $m_i(t)$ denote the number of times product i has been used up to and including time t . Determine whether the popularity of products follows a long tail. What does the exponent e tend to as k tends to infinity? [20pts]
4. The file `movies-hw2.txt` is an extract from recently published netflix data. The file contains the number of times the i -th movie was watched.
 - (a) Determine whether this data represents a long tail by finding the best fit of the form $a(b + i)^{-e}$. Explain your reasoning. [10pts]

- (b) Compute the median rank, the expected rank, and the standard deviation of a movie watched by a customer. Which of these would you expect to go to infinity as the number of movies increases but the tail exponent e remains fixed? [10pts]
- (c) Assume that netflix incurs a cost of \$0.90 for shipping and handling of each movie. Assume that a typical consumer rents 8 movies per month. Assume that the top 100 movies have a royalty fee of \$2 that netflix must pay, the next 900 have a royalty of \$1.50, and the rest have a royalty of \$0.50. Compute Netflix's cost per customer. How sensitive is this number to a \$0.25 change in the royalty fee for the first tier? For the third? [10pts]