FINAL EXAMINATION
Networked Life (NETS 112)
December 20, 2017
Prof. Michael Kearns

This is a closed-book exam. You should have no material on your desk other than the exam itself and a pencil or pen. If you run out of room on a page, you may use the back, but be sure to indicate you have done so.

Name:

Problem 1: ________/10

Problem 2: ________/10

Problem 3: ________/15

Problem 4: ________/10

Problem 5: ________/10

Problem 6: ________/10

Problem 7: ________/15

Problem 8: ________/10

Problem 9: ________/10

TOTAL: _________/100

Problem 1 (10 points). Answer "true" or "false" to each of the following assertions.

(a) In the game-theoretic model of packet or traffic routing we examined, there are examples in which the average latency at Nash equilibrium is twice as long as the optimal social welfare solution.

(b) In the behavioral experiments on biased voting, there were frequent instances of "spoilers" who prevented anyone from receiving any payoff.

(c) At any given moment of time, the IP address of a mobile device will (almost) always be unique across the entire Internet.

(d) In the U.S., there are laws and regulations designed to prevent the use of race as an input to credit scoring models.

(e) In the U.S., there are laws and regulations designed to prevent the use of social media activity as an input to credit scoring models.

(f) In the Travers and Milgram and Columbia Small Worlds experiments, roughly half of the initiated chains were completed.

(g) In the "Schelling diagrams" we used to model the evolving participation level in some group activity, it is possible for there to be any number of equilibria or stable states.

(h) Any game must have an odd number of Nash equilibria.

(i) The "long tail" of web search refers to queries that are individually rare but collectively frequent.

(j) The "Where's George" exponent was 2.0.

Problem 2 (10 points). Suppose a government agency is given no information about you other than the sequence of IP addresses that your mobile phone is assigned during a typical week for you at Penn. Do you think the agency could use that data to identify you? If your answer is "no", explain why not. If your answer is "yes", describe how the agency might do it.

Problem 3 (15 points).

(a) Write the payoff matrix of a 2-player, 2-action (pure strategy) game in which there is exactly one pure Nash equilibrium and no mixed equilibrium, and all payoffs are either +1 or -1. Specify the equilibrium.

(b) Write the payoff matrix of a 2-player, 2-action game in which there is exactly one mixed Nash equilibrium and no pure equilibrium, and all payoffs are either +1 or -1. Specify the equilibrium.

(c) Write the payoff matrix of a 2-player, 3-action game in which there is exactly one pure Nash equilibrium and exactly one mixed equilibrium, and all payoffs are either +1 or -1. Specify the equilibria.

Problem 4 (10 points). Pick any of the assigned scientific *articles* (not books) in which the main results relied on the collection, analysis or modeling of user or experimental data, and discuss potential privacy and/or fairness concerns with the methodology and/or published results and models. Your discussion should draw from the material and themes in "Data and Goliath" and "Weapons of Math Destruction".

Problem 5 (10 points). Briefly but precisely describe Zipf's Law for the frequency of English words. Then discuss the implications of the law for search engines and the monetization of search queries.

Problem 6 (10 points). Consider Kleinberg's model for forming a network mixing local and long-distance connectivity, but now consider a modified navigation algorithm in which a vertex forwards the message to the neighbor nearest the target with probability 0.9, and to a randomly chosen neighbor with probability 0.1.

(a) Consider the "typical" or average-case behavior of this modified navigation algorithm. Discuss whether you think it will generally result in faster, slower, or unchanged navigation time compared to the original navigation algorithm. If you think the answer depends on the value of the long-distance edge distribution exponent, explain why.

(b) Consider the worst-case behavior of this modified navigation algorithm. Discuss whether you think it will generally result in faster, slower, or unchanged navigation time compared to the original navigation algorithm. If you think the answer depends on the value of the long-distance edge distribution exponent, explain why.

.

Problem 7 (15 points). In lecture we considered a strategic or game-theoretic network formation model based on connectivity.

(a) Briefly but precisely describe the key elements of that model, including the pure strategies available to the players and what their payoffs were.

(b) We generally considered the "outcome" of a probabilistic or randomized network formation model to be the macroscopic statistical properties associated with the distribution over networks (e.g. preferential attachment tends to generate heavy-tailed degree distributions). What would we consider the "outcome" of a strategic network formation model to be?

(c) From the perspectives of both mathematical analysis and the real world, briefly discuss some pros and cons of both randomized and strategic network formation models.

Problem 8 (10 points). Consider the problem of scheduling final exams for Penn's Fall 2017 semester in a way that ensures that *no student has more than one final exam at a time, but the overall exam period is as short as possible*. We would like to cast this problem as an instance of the graph coloring problem. For each of the following parts, circle the most appropriate answer. Each part is worth 2 points, and *you will earn an additional 2 points if you get all four parts correct.*

(a) There should be a vertex in the network for each:

student at Penn Fall 2017 course exam date and time course instructor

(b) There should be an edge in the network between pairs of:

students in the same class

courses at the same or overlapping meeting times

students and instructors in the same class

students and the courses they are taking

(c) The colors should correspond to:

exam dates and times (e.g. Dec 20 at 9AM)

Fall 2017 courses

Penn students

course meeting times

(d) Minimizing the number of colors used corresponds to:

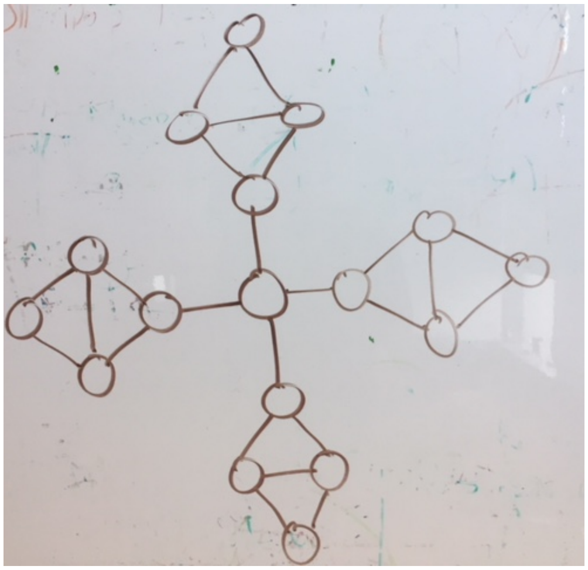
minimizing the number students with exam conflicts

minimizing the number of courses with the same exam times

minimizing the number of distinct exam slots needed

minimizing the number of course meeting times

Problem 9 (10 points). Consider the following network.



(a) Compute the clustering coefficient of the network.

(b) Compute the edge density of the network.

(c) Give the degree distribution of the network (number of vertices with each degree value).

(d) Give the worst-case diameter of the network.

(e) Give the chromatic number (minimum number of colors needed for solution to coloring problem) of the network.