CS5984: Urban Computing

Assignment 3

Assigned: Nov 16, 2015

Due date: Dec 03, 2015

1) **Computational Epidemiology (60 points):**

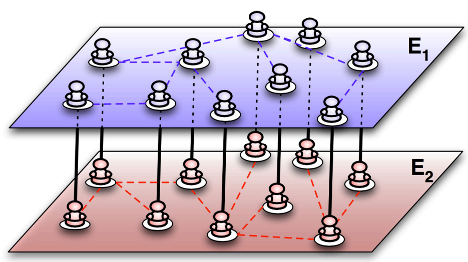
   a. ***Immunization Algorithms:*** Preventing the spread of a disease is an important problem in public health and urban computing. The first assignment asked one to design such immunization algorithms and the second assignment has given us a simulation framework to study disease transmission. In this question, we will evaluate how such immunization algorithms are implemented.

      i. Download the file at: https://www.dropbox.com/s/ichdkooh6dkozez/multiplex_networks.zip?dl=0

      ii. Run the SIR model till the system stabilizes (similar to the question in assignment 2) on *gowalla_edges_contact.txt*

      iii. Immunize (remove) nodes (and thus edges connecting them to the network) using any immunization algorithm of your choice e.g. friendship paradox. Re-run the SIR model on the immunized graph and observe the epi-curve. Evaluate and explain your results.

   b. ***Limiting disease spread:*** In real word complex systems, networks are composite/multiplex in nature. See figure below, for example:



   An individual can exist in several contact/communication networks or information networks. For instance, Alice can have accounts on Facebook and Twitter. While both these networks are different, the individual is the same. In this question, we will analyze how one contagion spreading in a network can act as a counter-contagion in another.

      i. Assume a variation of the SI model where a disease *I* is spreading on one network *N1* (*gowalla_edges_contact.txt*) and *F,* the fear about the disease, is propagating in the second network *N2* (*gowalla_edges_comm.txt*).

      ii. Start with k=10 nodes, β = 0.3 and simulate the spread of *I* for t=25 time steps. Plot the epi-curve of infection *I* at this juncture.

      iii. At time-step t = 26, introduce the second contagion *F* to the nodes and let it spread via *N2*, start with k = 10 nodes and β = 0.6. Choose these seed nodes strategically using a heuristic of your choice so as to maximize spread of fear *F*. Use this contagion (*F*) to design a *quarantining* algorithm – i.e. if a node is infected with *F,* it is removed from the contact network

N1. Plot the epi-curve for both infections. Does the epicurve of *I* affect the other?

iv. Does this quarantining approach limit the spread of the disease *I*? Evaluate the effectiveness of this strategy by simulating the disease spread without the quarantining algorithm and again with the quarantining algorithm. Explain the difference.

2) **Trajectory Mining (20 points):**
   a. Given a combination of trajectories of individuals via cellphone traces and interpersonal communications (via texts/Twitter/contact) how will you track the global spread of a disease? Assume that travel trajectories are known for individuals.
   b. If one had a large number of GPS based trajectories of taxicabs, how will you go about extracting temporal travel trajectories to assist in policy making for road transport networks?
   c. How can trajectory mining be applied to weather forecasting by tracking phenomena like hurricanes, tornadoes, and ocean currents?

3) **Energy disaggregation (20 points):**
   a. Using class slides and the two papers (http://www.mdpi.com/1424-8220/12/12/16838/htm and http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5735484) explain the differences between the supervised energy disaggregation and unsupervised energy disaggregation.
   b. Assume we have collected aggregated real power from a house and now we aim to separate two devices- a refrigerator and a dryer. How will you go about designing a disaggregation algorithm to tackle the above problem? Hint: think about how and when these devices are generally used in a household.