

RECONNECT 2015

Social Networks: Mathematical and Computational Tools with applications to Homeland Security June 14 - 20, 2015

Agenda

Sunday, June 14

- Check into hotel, Rochester NY
7:00 PM Dinner as a group in hotel

Monday, June 15

- 7:00-8:00 Breakfast in the hotel
8:15 Travel to meeting either by foot, car, or shuttle

All meetings are in Gosnell - A300 on the RIT campus. All lunches and breaks will be in the adjoining Atrium

- 8:45 AM Welcome, Introductions, and Plans for the Week
Midge Cozzens, Rutgers University, CCICADA and DIMACS Centers

Welcome from Rochester Institute of Technology Administration
President William Destler, Rochester Institute of Technology
Maurino Bautista, Professor of Mathematics, Campus Host

- 9:30 **Networks and their Applications to the Department of Homeland Security**
Dr. Raluca Gera, Naval Postgraduate School

This presentation builds on The Seven Bridges of Königsberg problem, the problem that is at the origin of graph theory, posed by Leonhard Euler in 1735. The talk will present a natural extension of this problem to today's type of problems that still need solutions: mapping large dynamical networks such as social networks, technological networks, networks of information and biological networks. We explore one particular technological network, the Internet, and share results comparing temporal snapshots of Internet's topology. We model the Internet by a graph, with router interfaces represented as vertices, and the connections between these interfaces -obtained through traceroutes- represented by the edges. The first part of the day is devoted to basic graph concepts and synthetic networks as models to understand and study the networks around us.

- 10:45 Break

- 11:00 Continue the discussion of graphs and modeling the internet using graphs.
- 12:15 Lunch in the Atrium
- 1:30 A challenging problem in studying large actual networks is that data about the network is generally incomplete: For most networks (such as Facebook or the Internet), it is impossible to observe all members of a network and all the interactions between them. Generally, researchers try to infer as much as possible of a network, and formulate it into a network such as the Internet as described above. An extension of our approach described above for the Internet is: create ground truth topology, infer that topology, and then compare the inferred to the true topology. We do this for four networks: two of them are models of graphs (an Erdős-Rényi graph and a Barabási-Albert graph used as baseline) and two real social networks (a Facebook and a Researchers collaboration network from Stanford's SNAP data set). We consider these networks being ground truth, and we infer them using two methodologies. In this talk we present our results of the node and edge count of the inferred networks compared to the true networks. We then propose to do the same for the Internet, but creating topologies that are Internet-like is not an easy task.
- 3:00 Break
- 3:15 Continue the discussion begun at 1:30
- 4:15 Work on homework and begin to think about module topics
- 6:00 Dinner at the hotel

Room available in the evening for work on homework and socializing

Tuesday, June 16

- 7:00-8:00 Breakfast in the hotel
- 8:15 Travel to meeting either by foot, car, or shuttle
- 8:45 AM Presentation of Homework Solutions
- 9:45 **Optimizing Disaster Response with Social Media**, Dr. Christie Nelson and Dr. Brian Nakamura, Rutgers University CCICADA Center
- In real-time emergency scenarios an accurate picture of the situation is needed quickly. Often during large-scale disasters, cell towers become overloaded, and the only way of communication is through text messages. It becomes important to gather information from text messages sent to emergency numbers in order to respond quickly and efficiently with life-saving efforts. In addition, responders are unable to manually handle the

large volume of incoming texts. Real-time information from streaming data is needed, and responders would benefit from text classification of incoming messages. To add to this difficult problem, these data sources tend to be microtext, which makes the problem of modeling the data more challenging. A discussion of a methodology to summarize text messages sent during an emergency used by responders, including the analysis of locations to identify geospatially potentially new areas of population in need of emergency assistance.

- 10:45 Break
- 11:00 Real-time disaster needs are input into a mixed integer programming resource allocation model for distribution of resources for disaster aid.
- 12:15 Lunch
- 1:30 Christie and Brian provide some hands-on activities. – **Be sure to have your laptop.**
- 3:00 Break
- 3:30 Two methods are discussed for determining emergency needs: A supervised method modeled the data with a variation of Naïve Bayes, Higher-Order Naïve Bayes (HONB), and an unsupervised approach modeled the data with a variation of Latent Dirichlet Allocation, Higher-Order Latent Dirichlet Allocation (HO-LDA).
- 4:00 Wrap up, homework assignment and suggested topics for modules or technical reports
- 5:00 Begin work on exercises and think about today's talks in light of modules; to be continued after dinner in groups
- 6:00 Dinner at the hotel

Wednesday, June 17

- 7:00-8:00 Breakfast in the hotel
- 8:15 Travel to meeting either by foot, car, or shuttle
- 8:45 AM Presentation of Homework solutions
- 9:45 **Developing and Testing New Analytic Tools for Discovering and Visualizing Social Media Data and Online Social Networks**
Dr. Anatoliy Gruzd, Ted Rogers School of Management and Director of the Social Media Lab at Ryerson University in Canada

As social creatures, our online lives just like our offline lives are intertwined with others within a wide variety of social networks. Each retweet on Twitter, comment on a blog or link to a Youtube video explicitly or implicitly connects one online participant to another and contributes to the formation of various information and social networks. Once discovered, these networks can provide researchers with an effective mechanism for identifying and studying collaborative processes within any online community. However, collecting information about online networks using traditional methods such as surveys can be very time consuming and expensive. Dr. Gruzd explores how the advent of social media and the growing availability of user-generated big data are changing the ways in which people communicate, collaborate and disseminate information and how these changes impact the social, economic and political norms and structures of modern society.

- 10:45 Break
- 11:00 The discussion begun earlier continues by exploring automated ways to discover and analyze communication networks from social media data.
- 12:15 Lunch
- 1:30 Participants learn how to use Netlytic (<http://netlytic.org>), a cloud-based text and social networks analyzer to collect, analyze and visualize publicly available online conversations from social media websites such as Facebook, Twitter and Instagram. Participants **are encouraged** to bring their laptops to follow the tutorial.
- 3:00 Break
- 3:30 Homework assignment and suggested topics for modules or technical reports
- 4:30 Begin work on exercises and think about today's talks in light of modules; to be continued after dinner in groups
- 6:00 Dinner at the hotel

Thursday, June 18

- 7:00-8:00 Breakfast in the hotel
- 8:15 Travel to meeting either by foot, car, or shuttle
- 8:45 AM Presentation of Homework solutions
- 9:45 Continue using Netlytic to discover and analyze online communication networks from social networking platforms such as

Youtube, Instagram and Facebook. (The first day will primarily focus on networks derived from Twitter.)

- 10:45 Break
- 11:00 Continue previous discussion
- 12:15 Lunch
- 1:30 Continue morning discussion
- 3:00 Break
- 3:30 Joyati Debnath will discuss modules in detail
- 4:00 Work on exercises and modules
- 6:00 Dinner at the hotel

Friday, June 19

- 7:00-8:00 Breakfast in the hotel
- 9:00AM Leave for tour of Toy Museum
- 9:30-12:00 Tour of Museum
- 12:30 Lunch at RIT
- 2:30 First module presentations - Joyati Debnath organizes
- 6:00 Dinner at the hotel

Saturday, June 20

- 7:00-8:00 Breakfast in the hotel
- 8:15 Travel to meeting either by foot, car, or shuttle
- 8:45-11:30 Group presentations - Joyati Debnath organizes
- 12:00 Box lunches and departure