# USC EE 599 – Social Network Systems – Fall 2014

**********  DRAFT April 30, 2014  **********

*Course Summary*

A social network is a group of people connected by some means of interaction. As such, social networks are as old as society itself and have been studied extensively in the social science literature. There are basic properties of social networks, such as polarization, which can be predicted by the strengths of the connections between individuals. One may also consider the dynamics of social networks  fashion trends, memes, viral videos, etc. While there are a few examples of very large social networks that have been extensively studied in the past (e.g., economics), many of the classical studies of social networks have been limited in scope because of the inability to conduct large-scale measurements and experiments.

Today, electronic social networks  both explicit (e.g., Facebook, Twitter, etc.) and implicit (e.g., corporate calendars, email traces, cell phone meta-data, etc.)  make it possible to systematically study these large-scale social networks. In fact, our world is changing rapidly. Data is constantly being collected on each of us and is used to influence a myriad of our decisions including purchasing and voting, and to garner our support for any number of causes. How these networks are structured and managed can have impacts on our life in unexpected ways from increasing our chances of divorcing to our religious choices.

The goal of this course is to bring a systems perspective to this new landscape. At a high level we would like to address the following questions:

- What are the general system properties of large-scale social networks?
- What causes a meme or epidemic to spread?
- How can large data sets be used to measure key properties of social networks?
- What mathematical tools can be used to translate these big data measurements to predictions on the social network?
- Can an entity substantially influence the behavior of a social network  e.g., Can planted Twitter users topple a regime? Can corporate leadership be computer automated?
- What are the requirements of the data center and network architectures that support this new world?

To address these issues, this course will survey the areas of Network Science, Data Analytics, and Data Center Architectures. I will also provide a review of mathematical basics (probability, statistics, linear algebra) and an introduction to the associated mathematical tools of the field (e.g., machine learning, game theory, optimization with sparse data, graphical models).

| | |
|---|---|
| **Instructor:** | Keith M. Chugg<br>EEB 500A<br>213-740-7294 (Voice); 213-740-8729 (Fax)<br>chugg@usc.edu – Include 599 in subject |
| **Office Hours:** | TBD |
| **TA:** | TBD |
| **Grader:** | TBD |
| **Lecture:** | Monday, Wednesday 12:00-1:20 in TBD<br>with 1:20-1:50 for advanced topics, office hours, discussion in the room. |
| **Webpage:** | Piazza Class Page for everything except grades<br>and (non-DEN) USC Blackboard Class Page for grades<br><br>– All HWs, handouts, solutions will be posted in PDF format<br>– *Student has the responsibility to stay current with webpage material* |
| **Prereqs:** | Graduate standing EE or CS |
| **Other Requirements:** | Basic computer skills (i.e., ability to program in some high-level language). |
| **Grading:** | 20% Homework<br>10% Participation (in-class and online)<br>35% Project 1<br>35% Project 2 |

**Exam Dates:** No exams. Projects and homework only.

**Projects:** Both projects will be team-based with no repeat teams. The scope and format of the projects will be determined by the enrollment. Teams will investigate areas of interest in more depth according to their backgrounds and interest (e.g., good with mathematics, good with programming, etc.). Some presentations will be required. Written reports or online modules will be required.

**Course Objective:** This is a survey class. The object is to gain a basic familiarity with the field of social network systems with some deeper level of exposure to selected areas through the class projects.

**Guest Lectures:** There will be guest lectures for 2-4 of the sessions by subject area experts.

**Grading Policies:**

- **Final grades** will be assigned by a combination of student score distribution (curve) and the discretion of the instructor.
- **Late HW** will not be accepted. A late assignment results in a zero grade.
- **Attendance:** Lecture attendance is encouraged but not mandatory. However, students are responsible for all material presented in lecture.

- **Statement on Academic Integrity:**

  USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect ones own academic work from misuse by others as well as to avoid using anothers work as ones own. All students are expected to understand and abide by these principles. Scampus, the Student Guidebook, contains the Student Conduct Code in Section 11.00, while the recommended sanctions are located in Appendix A. Students will be referred to the Office of Student Judicial Affairs and Community Standards for further review, should there be any suspicion of academic dishonesty. The Review process can be found at http://www.usc.edu/student-affairs/SJACS/.

- **Statement for Students with Disabilities:**

  Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to TA) as early in the semester as possible. DSP is located in STU 301 and is open 8:30 a.m.–5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776.

**Books and Reference Papers:**

- **Primary References:**

  1. D. Easley and J. Kleinberg, Networks,Crowds, and Markets: Reasoning About a Highly Connected World, Cambridge University Press, 2010. Available as pdf.
  2. A.-L. Baraba'si, Network Science. Available as pdf and iBook.
  3. Matthew O. Jackson, Social and Economic Networks, Princeton University Press, 2008, ISBN: 9781400833993. Several chapters available online.
  4. Jimmy Lin and Chris Dyer, Data-Intensive Text Processing with MapReduce, Morgan & Claypool Publishers, 2010. Available as pdf.

- **Secondary References (to be expanded):**

  1. Laszlo Lovasz, Large Networks and Graph Limits, Colloquim Publications, 2012, ISBN: 978-0821890851.

## Course Outline

Examples and applications will be given throughout the course.

1. Course Overview [1 week]
   (a) Types of social networks in modern life
   (b) Trends in big data and analytics
   (c) Application areas
   (d) Computation Social Science
   (e) Control and optimization of complex networks
   (f) Overview of research topics and trends
2. Review of Mathematics [1.5 weeks]
   (a) Probability and statistics
   (b) Linear Algebra
3. Graphical Models of Social Networks [1.5 weeks]
   (a) Graph properties and measures
   (b) Tie strengths and bridges
   (c) Community formation, polarization
4. Introduction to Mathematical Tools I [1.5 weeks]
   (a) Game Theory
   (b) Inference on graphs
5. Project presentations [0.5 weeks]
6. Network Dynamics [1.5 weeks]
   (a) Cascade effects in social networks
   (b) Power law models
   (c) Small world phenomena
   (d) Epedemics
7. Introduction to Mathematical Tools II [1.5 weeks]
   (a) Machine Learning
   (b) Optimization for sparse data sets
8. Data Analytics [1.5 weeks]
   (a) Map-Reduce
   (b) Distributed storage systems
9. Data Center Architectures: Trends and Challenges [1 week]

   (a) Historical overview
   (b) Connectivity challenges
   (c) Energy and economic considerations
   (d) Advanced memory architectures

10. Community Identification [1.5 weeks]
   (a) Clique detection in random graphs
   (b) Identifying influential members

11. Control of complex networks [1.5 weeks]
   (a) Predicting consensus
   (b) Tuning networks for stability/instability
   (c) Detecting anomalies

12. Project presentations [0.5 weeks]