CS 5984 - Social Computing

Fall 2018

Course Description

Social computing is a research area that is at the intersection of computational systems and social behavior. This course is geared toward developing a broad understanding of today's online social systems. From Twitter to Facebook and all the way back to email, social computing is one of the biggest forces on the internet. In this class, we will explore how and why social computing works? What are the real-world challenges and opportunities in current social computing systems? What's the right way to design these systems? What can you infer from the vast amounts of data people leave behind in these systems? What's still out there to build? What are some new emerging phenomenon that you are observing now in these systems? How is it changing our information ecosystem and what could be done about it? Specific course activities will broadly involve 1) applying quantitative methodologies to investigate and model data left behind in these systems to infer behavior and/or phenomenon, and 2) designing and building social tools that can augment current social computing systems so as to study a behavior in question or deploy small-scale experiments to test interventions in these systems. Aligned with best industry practices, students will be expected to work in a fast-paced, collaborative environment and to demonstrate independence and leadership.

Course Logistics

Instructor: Dr. Tanushree (Tanu) Mitra | tmitra@vt.edu | **Office**: Torg 3160E
Class Hours: Tues, Thurs 12: 30pm to 1:45pm | McB 226
Latest class schedule will be posted at: http://people.cs.vt.edu/tmitra/cs5984

Learning Objectives

After successful completion of this course, you will be able to:

- Identify important features of social computing.
- Assess the research issues in this field.
- Analyze data left behind in social media to answer questions from a variety of practical scenarios and domains, spanning politics, news, and health.
- Build social tools that augment current social computing systems.

Topics Covered

Below is a list of topics that will be covered in this class. The list is not exhaustive:

- Information disorders on social media (examples include misinformation, state-sponsored disinformation campaigns, astroturfing)
- Censorship online
- Algorithmic awareness, visibility, fairness, misinformation on web and social systems
- Designing systems that support online social processes (social translucence)
- Social networks (strength of ties, social capital in online networks)
- Online cooperation, identity, reputation, and deviant behavior (harassment, trolling)

Prerequisites

Graduate student standing. Prior background in some high level programming language is assumed. In terms of the required skills, students need to have basic knowledge of statistics, preliminary machine learning and some ease in implementing low-fi software demos. An overview of the concepts and tools needed will be reviewed in class, however in-depth coverage of the fundamentals is not in the scope of this course. Hence, students also need to be proficient in programming. Experience in use of a scientific computing software like R is a bonus, but not required. You are expected to quickly learn many new things. Below is a list of examples (not exhaustive) that your assignments and project activities may require you to do:

- fetch social media data using existing API (tweepy for Twitter, PRAW for Reddit)
- or fetch data via crawling and scraping (python scrapy, Beautiful Soup, Newspaper libraries),
- analyze posts from data using pre-existing libraries (like python *nltk, pandas*),
- modify an existing social media site (e.g., changing affordances for Facebook comments using Greasemonkey scripts, or by building a chrome plugin)
- or even design a new site or online tool to implement design principles behind these social computing systems or solve an existing problem faced by these systems (e.g. problems related to information disorders).

The associated programming basics will not be covered in class. Please make sure you are comfortable with this.

Assessment

Assessments will be conducted as a combination of the following components (*grade distribution subject to change*):

- Class participation – 10%
- Reading responses – 20%
- Assignments – 20 %
- Term project – 50%
    - Project Pitch – 5%
    - Project Proposal – 5%
    - Midterm Project presentation – 5%
    - Midterm milestone project report – 10%
    - Final project presentation/demo – 10%
    - Final report – 15%

Your work will be graded on a list of criteria (specified on the assignment) such as quality of writing, completeness, insight into design issues, insight into social issues, insight into technical issues, thoughts about ethical issues, etc. For each criterion, you will receive either a check plus, check, or check minus. Most criterion will receive a check. A plus means "you impressed me." A minus means the assignment is incomplete, incorrect, or sloppy in some fashion with respect to that criterion. Pluses and minuses are combined to give your grade for the assignment. For most assignments, you start out half way between a B+ and A-. One plus makes it an A-; one minus makes it a B+. These are general guidelines to let you know what to expect. Grading on specific assignments may differ.

Texts/Reference Materials

Since, social computing is an emerging area, the primary reading material for this course will be drawn from research papers. All daily readings will be provided as linked pdfs or as electronic reserve from the library website, and will include references such as (not an exhaustive list):

- Cheng, Justin, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. "**Antisocial Behavior** in Online Discussion Communities." *In Ninth International AAAI Conference on Web and Social Media*, (ICWSM) 2015.
- Erickson, Thomas, and Wendy A. Kellogg. "**Social translucence**: an approach to **designing** systems that support social processes." *ACM transactions on computer-human interaction (TOCHI)*, 2000.
- Lewandowsky, Stephan, et al. "**Misinformation** and its correction: Continued influence and successful debiasing." *Psychological Science in the Public Interest* 13.3 (2012): 106-131.
- Eslami, Motahhare, et al. "I always assumed that I wasn't really that close to [her]: Reasoning about **Invisible Algorithms** in News Feeds." *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. (CHI), 2015.
- Hannak, Aniko, et al. "Measuring **personalization** of web search." *Proceedings of the 22nd international conference on World Wide Web*. (WWW), 2013.
- King, Gary, Jennifer Pan, and Margaret E. Roberts. "Reverse-engineering **censorship** in China: Randomized experimentation and participant observation." *Science* (2014).
- Mitra, Tanushree, Graham P. Wright, and Eric Gilbert. "A parsimonious language model of social media credibility across disparate events." *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. (CSCW), 2017

Other highly recommended books (not required):

- Gillespie, Tarleton. *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press, 2018.
- Kraut, R. E. & Resnick, P. (2012). Building successful online communities: Evidence-based social design. Cambridge, MA: MIT Press. A third party summary of the book is available at https://acawiki.org/Building_successful_online_communities:_Evidence-based_social_design
- Olson, J. S., & Kellogg, W. A. (Eds.). (2014). Ways of Knowing in HCI (Vol. 2). New York, NY, USA:: Springer.
- Safiya Umoja Noble (2018). Algorithms of Oppression: How Search Engines Reinforce Racism