CS 4803 Social Computing: Introduction

Munmun De Choudhury

munmund@gatech.edu

Week 1 | August 17, 2014

Part I: Course Structure and Information

Learning Objectives

- Social computing is one of the biggest forces on the internet
- **Goal: How and why social computing works?**
  - What's the right way to design social computing systems? What sense can you make out of all the data people leave behind? What's still out there to infer, understand, and build?
- At the completion of this course, you'll be able to:
  - Understand important features of social computing
  - How to design and prototype new social computing systems
  - How to analyze data left behind in social media
  - Understand theoretical and practical research issues in this field

Grading

- Responses to Class Readings (on Piazza) - 20%
- Assignment I - 15%
- Assignment II - 15%
- In-class Presentation of Data Project Phase I (midterm) - 5%
- In-class Presentation of Data Project Phase II (final) - 5%
- Term project report and deliverables (due at final) – 30%
- Class Participation – 10%

Participation in Class Readings

- Readings are available on class website
  - Check the class webpage before each class, as readings are subject to change
  - http://www.munmund.net/CS4803.html
- Write short blurbs on Piazza (under "reading_reflections") about the readings assigned for a particular class
  - Blurbs can range from 200-500 words in length
- Blurbs should focus on the following, but not limited to:
  - The novelty in the paper you read i.e., what was the contribution
  - Why the particular method/data used in the paper was appropriate

Piazza

- Sign up with your GT information at this link [IMPORTANT]:

  piazza.com/gatech/fall2015/cs4803

- Class link:

  piazza.com/gatech/fall2015/cs4803/home

Assignment I

- Due: October 7, 2015 (3:35pm)
- Questions around the reading materials covered until the point the assignment is released.
- One design-focused question that would require submitting a mockup.

- What to hand in?
  - Write up of answers to questions in a pdf document, along with the mockup.
  - Submission on T-Square

Assignment II

- Due: November 25, 2015 (3:35pm)
- Questions around the reading materials covered until the point the assignment is released.
- One question that would test your thinking and formulation of applying data analytic approaches to a social computing question

  OR
- Implement data analysis on a small dataset that will be provided

- What to hand in?
  - Write up of answers to questions in a pdf document; if data analysis question is chosen, then code and a short summary of findings need to be submitted together in a zipped folder
  - Submission on T-Square

Term Project

- Project idea proposal due: September 30, 2015 (T-Square, 3:35pm)
- Mid-term presentations (phase I) due: October 28, 2015
- Final presentations (phase II) due: December 7, 2015
- Final project report and materials due: December 9, 2015 (T-Square, 3:35pm)
- Goals:
  - Group project: 2-3 people
  - You are free to pick your group
  - Need to discuss your project idea with instructor early on in the course. Instructor will help you frame and provide guidance on the project
  - Presentations (both midterm and final) will be total 20 minutes, with 15 minutes talking about task accomplished, and rest for Q&A

What to hand in?

- If building a tool: working prototype, and proposal of a method of its evaluation
- If analyzing data: choice of the analysis technique (text vs. network analysis), key findings from the analysis (e.g., charts and tables)
- If literature review or a research proposal: ~10 page long report, single spaced, single column format (excluding figures, tables and references)
- Notes:
  - Clearly articulate in an extra page individual contribution
  - Okay to use open source tools (in fact encouraged)

Finding Data

- Based on your term project, if you need data…
- Suggestions:
  - Your own social media data. E.g., you might consider downloading all of your email or IM logs. Or, you could download all your Facebook or Twitter data.
- Hillary Mason's research dataset page: https://bitly.com/bundles/hmason/1
- Jure Leskovec's SNAP page: https://snap.stanford.edu/data/

Late Policy

- Reading responses are due at 11:59pm on the day before the relevant class meeting.
- Assignments and project deliverables are due *one hour before class* (i.e., 3:35pm) on the date listed for that assignment/ project.
- Any work (reading response, assignment, or project) submitted more than 15 minutes after the due time will be assessed a 10% penalty.
- Each additional 24 hours of lateness will result in an additional 10% being taken off the grade for that assignment. After 5 days, the assignment will not be accepted and a grade of 0% will be entered.

English as Second Language

- If English is not your first language, you may request to not be graded on your writing for a particular individual assignment.
  - This means you won't be penalized for bad writing, but you also won't get credit for good writing. To take advantage of this option, you must mark "ESL" (English as a Second Language) on the first page of your assignment/paper.
- This option is not available for the term project as it is a group assignment.

Academic Integrity

- This class abides by the Georgia Tech Honor Code.
- All assigned work is expected to be individual, except where explicitly indicated otherwise.
- You are encouraged to discuss the assignments with your classmates; however, what you hand in should be your own work.
  - Okay to use open-source software (no need to reinvent the wheel), however do acknowledge!
  - Copying/reusing code from your classmates and friends are not allowed; strict action will be taken if similarities are discovered
  - Copying (textual) content for your assignments and project from other published work (without citing them) is also not allowed, and is considered plagiarism

Required Skills

- Technical: any object-oriented/scripting language like Python, Perl, C#; some frontend development skills/ web programming skills (ajax, javascript, php) if your project is about building a social tool

- Reading/writing: significant weightage on weekly class readings, mid-term and final term project report

- *Bonus:* Statistics/data mining: preliminary knowledge of working with some data and using some analytical software

No required text, only suggestions

- Social Network Analysis:
  - Networks, Crowds, and Markets, by David Easley and Jon Kleinberg
  - Six Degrees, by Duncan Watts
- Social Science-y:
  - On Individuality and Social Forms, by Georg Simmel
  - Networked, by Barry Wellman
- Machine Learning:
  - Machine Learning for Hackers, by Drew Conway and John Myles White
  - Natural Language Processing with Python, by Steven Bird, Ewan Klein, and Edward Loper
  - Pattern Classification, by Richard Duda, Peter Hart, and David Stork
- Pop Statistics:
  - The Signal and the Noise, by Nate Silver
- Writing:
  - Writing for Social Scientists, by Howard Becker

Help and Resources

- Office hours: 11am – 12 noon Fridays, or by appointment
- Location: TSRB 341

- Teaching Assistant: Trustin Clear
- Office hours: TBA
- Location: TBA
- Email: trustin@ gatech.edu

- Class website (including readings): http://www.munmund.net/CS8803.html

General Resources

- Sample reading reflections will be made available on Piazza, under Resources
- Assignments will be released three weeks ahead on Piazza, under Resources
- Class slides (already lectured) can be obtained from the class website:

  http://www.munmund.net/CS4803.html

- Important announcements will be made over email (via T-Square)

Part II: Introductions

- name + program
- background
  (e.g., tech/design)
- what you want to learn from the class

Part III: Defining “Social Computing”

Defining Social Computing

The Social Life of Small Urban Spaces by William H. Whyte:

https://vimeo.com/111488563

Syllabus and topics to be covered

- Syllabus: http://bit.ly/1KrXtma
- Key characteristics
  - SNS, social media overview; design of social systems
  - Discourse and dialogue; social capital
- Methods
  - Text analysis; network analysis
- Key issues of social computing
  - Credibility; trust; polarization; privacy
- Uses of social computing systems
  - Personality; trends and forecasting; event analytics; social news; social search and Q&A; location and mobility