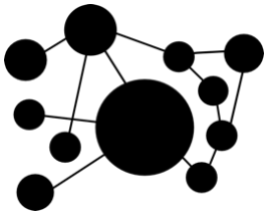

Social Network Analysis
COMS930
FA2019

Instructor: Cameron W. Piercy, PhD.
Email: cpiercy@ku.edu
Office hours: T: 10 – noon, W: 10 – 11a, and by appointment:
tinyurl.com/piercycal & tinyurl.com/zoomwithpiercy
Office: Bailey Hall 6C
Course time/location: Monday's 3:30 – 6:30pm , BA401

Course Overview

This course reviews theoretical, conceptual, and analytic issues associated with network perspectives on communicating and organizing. The course will review scholarship on the science of networks in communication, economics, organizational science, political science, psychology, and sociology. In the process we will take an in-depth look at theories, methods, and tools to examine the structure and dynamics of networks.

The majority of class time will be spent discussing the assigned readings and practicing network analysis in *R*. Students will be expected to complete small analysis assignments each week. Students will practice computer-based network analysis, modeling, and visualization tools. Students will either individual write a term paper reviewing literature in an area of interest related to social networks and proposing a research design assuming full access to the ideal dataset OR complete an empirical team research project suitable for conference submission.

Course Learning Goals and Objectives

By the end of this course, you will be able to:

- Use network theories for your own research;
- Use network analysis as a research technique (e.g., gather the data, identify relevant concepts, and analyze and represent the data);
- Explain how network concepts apply to theories and topics which interest you;
- Analyze relationships in terms of social networks, and organization theory;
- Synthesize networks theories and concepts,
- Use computer software to manage raw network data and then conduct various analyses.

Course Format

This is a graduate level course comprised primarily of PhD students. The course is discussion based, and all students are expected to have completed the class reading prior to the assigned day. Classes will feature a minimal amount of direct lecture and PowerPoint, and will focus on discussion of key ideas in the articles assigned for a given class and practice running the software. We will mix theory and practice, and students are encouraged to bring their own experiences/data/theories into the classroom.

Expectations, Requirements and Grading

There are three requirements for the course: UseR labs, one term paper, and in-class participation (including article discussion leadership). Each chapter of the Luke (2015) text involves data manipulation in the R environment. Students will conduct their own analyses on network data using either exercise problems in the textbook or techniques that mirror practices in the articles we read. Equal emphasis will be given to conducting the analysis and interpreting (and reporting) the results..

Writing is a critical part of communication, and in order to be successful in this class, it is important that you communicate clearly and concisely in writing. KU offers a Writing Ceneter where students can obtain help with writing skills and assignments: http://writing.ku.edu/ku-graduate-students. Students are encouraged to take advantage of their tutoring services before handing in any written work.

Software and SNA

Many texts contend that UCINET is the most commonly used software for network analysis (e.g., Borgatti et al., 2018). However, as the class-readings will demonstrate, R has become the gold-standard for both visualizing and analyzing network data. Thus, I have made the choice to use Luke's (2015) guide and require students to use R for data analysis. R is a free software with a dynamic set of packages capable of doing almost any computation available in all other software. An additional benefit is that R is not platform dependent.

I recommend students (even those experienced in the R environment) download the latest version of R: https://cran.r-project.org/

I would also strongly recommend students install the RStudio Environment to improve work-flow: https://www.rstudio.com/products/rstudio/#Desktop

> *Note:* Download and install R then download and install RStudio. These are free software you should not have to pay for anything.

This guide and video may be helpful in installation:
https://www.markanthonyhoffman.com/social_network_analysis/installing-r-and-rstudio.html
https://www.youtube.com/embed/FDSmlIBy7ko?rel=0

Your final grade will be allocated between the three requirements as follows:

- UseR Network Analysis Instructions: 40% (Submit 10 assignments across 11 weeks)
- Participation (Article discussion leadership and in-class discussion): 30% (20% + 10%)
- Term Paper: 30%
- Drafted paper (friendly) review: Uncompensated, like all reviews.

UseR Instruction Labs (40% total): Between September 2 and November 18th students are expected to turn in a ".R" file showing they tried a technique presented in one of the readings (e.g., centrality, visualization (of two-mode network, by attribute, using a particular layout, etc.), correspondence analysis, etc.) Students must turn in an assignment ten of the 11 weeks. Many

techniques are explicitly outlined in Luke (2015) and students are encouraged to modify the syntax provided by Luke (2015) for their own instructions. Any command not outlined in Luke's (2015) text is VERY likely to have a step-by-step guide online which you can model after. The file you turn in should include annotations using the "#" to explain relevant observations, notes, annotations, etc. associated with the syntax you generated. **Please label each exercise using the following convention: 'LastName_Technique_DATE.R' or 'Piercy_StructHoles_10.7.R'**

Students are encouraged to use the datasets included in the R Package 'UserNetR' which are described as follows:

	Net Type	Nodes	Edges
Bali	Whole, Directed One-Mode	17	126
Coevolve	Four iGraph Networks (fr_w1, fr_w2, fr_w3, fr_w4)	37	4 networks--varying
DHHS	Whole, Undirected, One Mode	54	477
FIFA_Nether	Whole, Directed, One Mode	11	108
Facebook	Whole, Undirected, One Mode	93	323
ICTS_G10	Whole, Undirected, One Mode	493	1359
Krebs	Whole, Undirected, One Mode	19	27
Moreno	Whole, Undirected, One Mode	33	46
PowerGrid	Whole, Undirected, One Mode	42	87
Simpsons	Whole, Undirected, One Mode	15	24
TCnetworks	Four iGraph Networks	25	4 Varying Networks
hwd	Two-mode network	1365 (160 movies, 1205 actors)	1600
Ihds	Whole, Undirected, One Mode	1283	2708
middleschool	Whole, Directed, One-Mode	37	179

If students wish to use a supplemental/alternative dataset they are welcome to do so. When using other datasets, please include a .csv or other accessible file format with your assignment submission.

These assignments are meant to (1) create a repository of syntax used to manipulate your own data and (2) help you work on real data using R. We will share our syntax in a folder to generate a class-library of techniques. Late assignments will be penalized 25% per day.

Participation & Discussion Leadership (30%; 15% for discussion leadership, 15% for presentations): Students will be responsible for leading discussion on given days. Students will serve as discussion leaders each class and will explain the readings and empirical studies. The goal is to put the readings in terms which help to generate discussion based on:

- The objectives of the studies and an overview of the topics covered.
- Theories addressed in the research
- Aspects of the design and implementation of the studies
- Key contributions of the study and questions/issues it provokes
- Ways in which the reading connects to other research
- A list of several discussion questions which remain unanswered in the papers

We will follow Amazon rules for discussion; each student leading discussion will be expected to provide a 1-page summary over the topic covered that day and list key questions (both theoretical and methodological). Bullets are encouraged. Your job is to identify key aspects of the readings explained in plain language (your own words, direct quotes should be minimized as much as possible), and to find common themes, discoveries, controversies, etc., across articles. Include an **opening comment on the connections among the articles. Upload the summary and questions for that day onto Blackboard Discussion Forum by the class time.**

Graded presentations are the visualization milestone and presentation of the final paper.

Final Term Paper (30%): Students may choose how to approach their term paper: (1) an individual rigorous proposal or proposition paper articulating relationships among network dynamics and existing theory or (2) a team-based empirical essay using existing data to generate new understandings.

In both cases, the term paper should develop or elaborate a theory, and explicitly incorporate the network perspective. The team (2 or 3) empirical paper should include methods, results, and a discussion. The individual paper should be ready for data collection or presentable as a proposition paper at a conference. The team empirical paper should be ready for conference submission. Either way, the paper should review relevant research literature and include testable network hypotheses or makes novel methodological or computational contributions. Papers need to be prepared according to the guidelines specified in the Publication Manual of the American Psychological Association (6th Ed.). You are free to use this as an opportunity to develop ideas you have worked on in other courses or for thesis/dissertation. If you are using an idea you've worked on in the past, inform Dr. Piercy BEFORE you begin the project: Avoid self-plagiarism.

The term paper is due on ***Wednesday, December 11th***. The whole paper should be complete (including any Figures, Tables, References, and Appendices), using 12-point font, double-spaced, and with 1" margins.

Students will present their term paper on the last day of the course. Presentations will be 10-15 minutes in length, including time for discussion/Q&A. The presentations are meant to be a forum for sharing the knowledge you've developed.

Students, especially those taking on the team project, are encouraged to consider the potential for a strong secondary analysis of existing datasets. Please spend significant time exploring potential datasets before you settle. There are many publicly available datasets and I want you to find one that fits your interest, needs, and skill. Here are some potential sources of data:

- https://toolbox.google.com/datasetsearch
- http://networkrepository.com
- https://www.redeftiedata.eu/

Team Empirical Paper: Your team will need to find a dataset to execute network analysis techniques. I want encourage you to find a network which matches your abilities and interests. If you have a lot of experience with R, you may wish to find a large network and attempt new

analysis techniques. If you are new to R, you will want a smaller network (i.e., > ~200 nodes with as few as 20-30) to feel confident in your findings. Here are some general rules-of-thumb:

- Choose a network that overlaps with your research interests
- Plan to write a paper about the composition of this network. Consider early if you are interested in a descriptive approach to the network, traditional multi-variate approaches to the data, or stochastic models.
- Please meet with Dr. Piercy early to discuss your network data.
- You may also wish to scrape your own data. I have limited knowledge about tools for scraping data. However, I hear the following tool can be quite useful. I'll provide any support I can if you wish to obtain your own data.
    - https://mkearney.github.io/nicar_tworkshop/#1

Grades: All grades are final. Please do not ask to have your grade changed for reasons other than mathematical error. Applying subjective standards after the fact invalidates the standards applied to the entire class and is unfair to every student.

An incomplete can only be assigned to a student who due to unforeseen (and generally emergency) circumstances cannot finish coursework within the given semester. Students should not assume that they have the option of an I grade; this option will be utilized infrequently and at the professor's discretion.

A note on communication:

• Emails for this class should always start with "SNA" as the beginning of the subject.
• Include a salutation
• Sign your emails with your full name.
• Clearly indicate what your question is. The more relevant information you provide, the more helpful and timely my response can be.
• Use full sentences and proper grammar.

University Policies

Religious Holidays: Students observing religious holidays that may require them to miss class periods or scheduled exams, speeches, assignment due dates, etc., should contact Dr. Piercy privately within the first two weeks of class to arrange alternative times for completing assignments and/or have these absences excused. If any schedule change interferes with a religious observance, please contact me as soon as possible for alternative arrangements.

Student accessibility and success: Any student needing accommodations for the course should let the instructor know. The earlier we coordinate any accommodations, the better I can support your learning. Students who need assistance obtaining accommodations may contact Student Access Services in 22 Strong Hall and can be reached at 785-864-4064 (V/TTY). Information about their services can be found at http://www.access.ku.edu. Please contact you instructor privately in regard to your needs in this course.

Mandatory Reporting: With very few exceptions, all employees at the University of Kansas are required to contact the Office of Institutional Opportunity and Access (IOA) at 785-864-6414 or ioa@ku.edu to report incidents of discrimination and sexual harassment, including sexual violence, of which they know or have reason to believe may have occurred. For example, if a student shares information about discrimination or sexual harassment, including sexual violence, with a faculty or staff member, the faculty or staff member must report the information to IOA. Asking faculty or staff members to keep something "confidential" does not exempt them from the mandatory reporting requirement.

Academic Misconduct: Academic misconduct is a serious offense. Academic misconduct is described in *Article II, Section* 6 of the University Senate Rules and Regulations. You are responsible for knowing the standards of academic conduct. The document is available here: policy.ku.edu/governance/USRR

Plagiarism: Plagiarism is a serious offense. Using the words and ideas of others is borrowing something from those individuals. It is always necessary to identify the original source of supporting information. **You must cite the source of any material, quoted or paraphrased, in both written work and oral presentations.**
Sometimes writers are uncertain about what to cite. Here are two firm guidelines:

- If you write word for word what appears in another source, put quotation marks around it and cite the source (author, year, page number).
- If you borrow and summarize ideas, arguments, data, or other information from another source, cite the source even if you put the material in your own words (author, year).
- Agreeing with the material does not make it your own; if it originated with someone else, give that person credit according to a formally recognized style.
  Helpful websites:
  - http://writing.ku.edu
  - https://owl.purdue.edu/

Course Readings

Required:

Luke, D. A. (2015). A UseR's Guide to Network Analysis in R. *Springer Publishing*.*
Available for free at:
http://www2.lib.ku.edu/login?URL=https://ebookcentral.proquest.com/lib/ku/detail.action?docID=4199226

American Psychological Association (2012). *Publication manual of the American Psychological Association* (6th Ed.). Washington, DC: APA.

Optional readings and resources:

https://github.com/briatte/awesome-network-analysis*
http://sna.stanford.edu/rlabs.php*
https://www.jstatsoft.org/issue/view/v024*
https://www.markanthonyhoffman.com/social_network_analysis/index.html*

http://www.shizukalab.com/toolkits*
https://kateto.net/networks-r-igraph*
https://docs.google.com/spreadsheets/d/1CoFGtrW85D9FsVcAE5-bcXVl6QOTncwXjFBYp4u2WgE/edit?usp=sharing* A spreadsheet containing nearly all R packages that deal with network analysis.

Barabasi, A. (2003). *Linked: How everything is connected to everything else and what it means*. London, England: Penguin Books.

Borgatti, S. P., Everett, M.G., & Johnson, J. C. (2018). *Analyzing social networks* (2nd ed.). London, UK: Sage.

Burt, R. S. (1992). *Structural holes.* Cambridge, MA: Harvard University Press.

Crossley, N., Bellotti, E., Edwards, G., Everett, M. G., Koskinen, J. & Tranmer, M. (2015). *Social Network Analysis for Ego-Nets.* London, UK: Sage.

Easley, D., & Kleinberg, J. (2010). *Networks, crowds and markets: Reasoning about a highly connected world.* New York: Cambridge. Available here: http://www.cs.cornell.edu/home/kleinber/networks-book/

Hanneman, R A., & Riddle, M. (2005). *Introduction to social network methods*. Riverside, CA: University of California, Riverside. Available here: http://faculty.ucr.edu/~hanneman/nettext/index.html

Hansen, D., Shneiderman, B., & Smith, M. A. (2010). *Analyzing social media networks with NodeXL: Insights from a connected world*. Burlington, MA: Elsevier.

Kadushin, C. (2012). *Understanding social networks: Theories, concepts, and findings.* New York: Oxford University Press.

Monge, P., & Contractor, N. S. (2003). *Theories of communication networks*. New York Oxford University Press.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and application.* Boston, MA: Cambridge University Press.

Watts, D. J. (2003). *Six degrees: The science of a connected age*. New York: W. W. Norton & Co.

*Resources with * are PARTICULARLY helpful for UseR Instructions Homework*

Software Packages:
R: R is powerful for general network analysis and visualization. The R platform is necessary for the statistical models of networks such as exponential random graph modeling (ERGM). The main packages in R include SNA, Statnet, ggplot, and iGraph. We'll explore many packages, but we will not use all of the available R options.

UCINET: Borgatti, S., Everett, M., & Freeman, L. (2005). UCINET 6 for Windows software for social network analysis. Harvard, MA: Analytic Technologies. www.analytictech.com. Free trial for 30 days is available and student discount price is $40 for a permanent purchase.

StatNet: Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., & Morris, M. (2007). Statnet: An R package for the statistical modeling of social networks. Funding support from NIH grants R01DA012831 and R01HD041877. http://www.csde.washington.edu/statnet

PNet: Wang, P., Robins, G., & Pattison, P. (2007). Software that includes procedures for MCMC MLE for exponential random graph models. University of Melbourne, Australia. http://www.sna.unimelb.edu.au/pnet/pnet.html

NodeXL: Smith, M., Milic-Frayling, N., Shneiderman, B., Mendes Rodrigues, E., Leskovec, J., Dunne, C. (2010). NodeXL: a free and open network overview, discovery and exploration add-in for Excel 2007/2010. Social Media Research Foundation. http://nodexl.codeplex.com/

SocioViz: Social Media (mostly Twitter) analytics platforms powered by Social Network Analysis metrics. https://socioviz.net/

Pajek: Batagelj, V. & Mrvar, A. (2010). Pajek–Program for Large Network Analysis. University of Ljubljana. http://mrvar.fdv.uni-lj.si/pajek/

Tentative Schedule

Subject to change based on class demands

August 26

What is the Network Perspective?

Wasserman & Faust, 1994, Ch. 1
Borgatti, Mehra, Brass, & Labianca, 2009
Freeman, 2008

September 2

No class: Labor Day

Luke, Ch. 1, 2, & ESPECIALLY 3
Submit first homework demonstrating that R is installed on your computer and you have completed three of the techniques outlined by Luke. *Individual meetings with Dr. Piercy during this week are encouraged.*

September 9 – Guest Lecture, Dr. Sun Kyong Lee, Associate Professor, University of Oklahoma

Types of Network Data

Ego-Networks: Crossley et al., 2015
Whole Networks: Kadushin, 2012, Ch. 3; Ch. 4;
Key terms: Borgatti, 1994
Optional: Ziberna, 2007 on R blockmodeling package

September 16

Bases for SNA

Monge & Contractor, 2003, Ch. 2
Borgatti & Lopez-Kidwell, 2011
Shumate, Pilny, Atouba, Kim, Peña-y-Lillo, Cooper, Sahagun, & Yang, 2013

September 23

Collecting and Organizing Network Data

Borgatti et al. 2018, Ch 4
Scott, 2017, Ch. 3
Kadushin, 2012 Ch. 11 Ethics
Eddington, 2019

September 30

Visualization

Luke, Ch. 4 & 5
Ognyanova, 2016 (https://kateto.net/netscix2016.html)

***Milestone assignment: Present your own visualization of some network data.

October 7

Basic Organizing Principles

- Homophily: McPherson et al., 2001
- Strength of weak ties: Piercy & Lee, under review
- Structural holes: Burt, 2005
- Network formation: Bakshy et al., 2015

October 14

No Class: Fall Break

Submit a rationale for your study, this document must explain why the topic you chose fills an existing research gap or extends our knowledge of network theory. Your rationale should be no longer than 5 pages and should be extensively cited (especially with sources beyond class readings).

October 21

Network Shapes and Formation

- Watts, 2004
- Klarreich, 2018 (https://www.quantamagazine.org/scant-evidence-of-power-laws-found-in-real-world-networks-20180215/)
- Lazer & Friedman, 2007
- Barabàsi & Bonaneau, 2003 (http://barabasi.com/f/124.pdf)

October 28

Network Contagion and Centrality

- Pilny & Prolux, 2019
- Luke, 2015, Ch. 7
- Su, 2012
- Lee & Katz, 2015

November 4

Social Capital, Structural Holes, and Grouping

- Appel et al., 2014 **OR** Burt, 2015
- Obstfeld, 2005
- Luke, 2015, Ch. 8, Ch. 9

November 11

Multivariate Statistics and Network Analysis

- Lutz & Hoffman, 2018
- Piercy & Lee, 2019
- Liang & Fu, 2019

November 18

Testing Hypotheses

- Lai, She, & Ye, 2019
- Siew & Vitevitch, 2019
- Saffer, 2016

November 25:

Submit complete draft of paper. This document should represent a nearly-finished product. Dr. Piercy will read each paper and provide detailed feedback to help facilitate a successful (i.e., conference ready) final product.

December 2

Network-Change Over Time

Kossinets & Watts, 2006
Seo & Thorson, 2015
Kearney, 2019

December 9

Stochastic Network Models

Luke Ch. 10, 11, & 12
Goodreau, Kitts, & Morris, 2009
Pilny & Atouba, 2018
Optional:
http://statnet.csde.washington.edu/workshops/SUNBELT/current/ergm/ergm_tutorial.pdf
Lee, Kim, & Piercy, 2019

December 16

*****Final presentations of individual and team papers *****