

Dynamic Network Analysis (Spring 2017)

Course 08-801, 08-640, 19-640

Instructor

Prof. Kathleen M Carley
Phone: x86016
Office: Wean 5130, ISR
E-Mail: kathleen.carley@cs.cmu.edu
Office Hours: Wednesday or by appointment

Teaching Assistant

Sumeet Khumar & Binxuan Huang

Introduction

Who knows who? Who knows what? Who communicates with whom? Who is influential? How do ideas, diseases, and technologies propagate through groups? How do social media, social, knowledge, and technology networks differ? How do these networks evolve? How do network constrain and enable behavior? How can a network be compromised or made resilient? What are network cascades? Such questions can be addressed using Network Science. Network Science, a.k.a. social network analysis, link analysis, geo-network analysis, and dynamic network analysis is a fast-growing interdisciplinary field aimed at understanding simple & high dimensional networks, from both a static and a dynamic perspective. Across an unlimited application space, graph theoretic, statistical, & simulation methodologies are used to reason about complex systems as networks.

An interdisciplinary perspective on network science is provided, with an emphasis on high-dimensional dynamic data. The fundamentals of network science, methods, theories, metrics & confidence estimation, constraints on data collection & bias, and key research findings & challenges are examined. Illustrative networks discussed include social media based (e.g., twitter), disaster response, organizational, semantic, political elite, crises, terror, & P2P networks. Critical procedures covered include: basic centralities and metrics, group and community detection, link inference, network change detection, comparative analytics, and big data techniques. Applications from business, science, art, medicine, forensics, social media & numerous other areas are explored. Key issues addressed: Conceptualization, measurement, comparison & evaluation of networks. Identification of influential nodes and hidden groups. Network emergence, evolution, change & destabilization.

In this course, the fundamentals of network science, the methods, the theories, the constraints on data collection are examined. This graduate seminar, offers an

overview and evaluation of the theory and research on networks broadly defined. Student are encouraged to bring and use their own data, or to use one of the large number of datasets available publicly in this area for assignments. Questions addressed include, but are not limited to: How do we conceptualize, measure, compare and evaluate various types of networks? How do we evaluate the impact of policies and technology on using these networks especially given the fact that these networks are dynamic? What nodes, relations, groups, motifs stand out in or are influential in a network? How do networks emerge, evolve, change? What is the difference in analyzing networks as complete graphs versus networks as emerging from a set of links? How can data on networks be collected and what are the limits of these collection techniques?

Prerequisite: Undergraduate-level statistics course or instructor permission. Linear algebra is recommended but not required. Students are encouraged to bring & use their own data, or to use provided data.

Course Content

Lecture slides and supplemental readings are available in the Course Content section of Blackboard. Weka data mining software is freely available and can be downloaded from <http://www.cs.waikato.ac.nz/ml/weka/downloading.html>.

Software

Required Software

ORA-LITE, available from CASOS -
<http://www.casos.cs.cmu.edu/projects/ora/>
Optional Software - R
ORA-PRO --- to be provided by Dr. Carley
NetMapper --- to be provided by Dr. Carley
Note - on all problem sets you must use ORA.

Available Software

AutoMap, available from CASOS -
<http://www.casos.cs.cmu.edu/projects/automap/>
Construct, available from CASOS -
<http://www.casos.cs.cmu.edu/projects/construct/>

Books

Required Books

Get a copy of Wasserman and Faust (SNA). Many books are available on-line. Papers are on Blackboard.

Wasserman, S. & K. Faust, 1994, *Social Network Analysis: Methods and Applications*. Cambridge University Press.

Carley, K.M. 2017. *Dynamic Network Analysis*.

Available at: http://www.casos.cs.cmu.edu/projects/book/DNA-Book_Draft.pdf

Recommended Books (to be aware of)

Marina Hennig, Ulrik Brandes, Jürgen Pfeffer, and Ines Mergel, 2014, *Studying Social Networks: A Guide to Empirical Research*, University of Chicago Press

Ian McCulloh, Helen Armstrong & Anthony Johnson, 2013, *Social Network Analysis with Applications*, Wiley

Sean Everton, 2012, *Disrupting Dark Networks*, Cambridge University Press

John Scott and Peter J Carrington, 2011, *The SAGE handbook of social network analysis*, Sage Publications

David Easley and Jon Kleinberg. 2010, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press.

National Research Council, 2006. *Network Science*

<http://www.nap.edu/catalog/11516.html> or

<http://www.nap.edu/books/0309100267/html/>

Mark Newman, D.J. Watts and A. Barabasi, 2006, *The Structure and Dynamics of Networks*, Princeton University Press.

Carrington PJ, Scott S, and S. Wasserman, 2005, *Models and Methods in Social Network Analysis*. Vol. 28. Cambridge university press

Ulrich Brandes and T. Erlebach, 2005, *Network analysis. Methodological Foundations*. Springer: Heidelberg (Germany).

Linton Freeman, 2004, *The Development of Social Network Analysis: A Study in the Sociology of Science*. Vancouver: Empirical Press.

Ronald Breiger, Kathleen M. Carley, and Philippa Pattison (Eds.). 2003.

Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers. Committee on Human Factors, Board on Behavioral, Cognitive, and Sensory Sciences. Washington, DC: National Academy Press.

Albert-László Barabási and Jennifer Frangos. 2014. *Linked: the new science of networks science of networks*. Basic Books.

Duncan J. Watts, 1999. *Small worlds: the dynamics of networks between order and randomness*. Princeton university press, 1999.

Duncan J. Watts, 2002, *Six Degrees: The Science of a Connected Age*, New York & London: W.W. Norton & Company.

Take care of yourself.

Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. You are not alone. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is often helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call 412-268-2922 and visit their website at <http://www.cmu.edu/counseling> Consider reaching out to a friend, faculty or family member you trust for help getting connected to the support that can help.

If you or someone you know is feeling suicidal or in danger of self-harm, call someone immediately, day or night:

CaPS: 412-268-2922

Re:solve Crisis Network: 888-796-8226

If the situation is life threatening, call the police:

On campus: CMU Police: 412-268-2323

Off campus: 911

If you have questions about this or your coursework, please contact Prof. Kathleen M. Carley kathleen.carley@cs.cmu.edu

Assignments, Grading, and Late Work Policy

Homework: 6 x 100 point Problem Sets

- **HW1**
- **HW2**
- **HW3**
- **HW4**
- **HW5**
- **HW6**

Course Project: 600 points

- **Project Proposal:** Due 2/20; A brief discussion of your proposed research problem, dataset, methods, and expected challenges. No more than 500 words.
- **Project Draft:** Due 3/20; Intro, Background, and Methodology Sections
- **Project Presentation:** To be scheduled in late April, 10-15 minute conference talk discussing your project's methods and key findings
- **Project Write Up:** Due 5/5; Publication style write up of your project. See examples on Blackboard

Late Work Policy: You are expected to turn in all work on time. Because we understand that exceptional circumstances may arise, each student will be permitted to turn in two of their assignments up to 48 hours late with no penalty. Any other late assignments will not be accepted.

For assistance with the written or oral communication assignments in this class, visit the Global Communication Center (GCC). GCC tutors can provide instruction on a range of communication topics and can help you improve your papers and presentations. The GCC is a free service, open to all students, and located in Hunt library. You can make tutoring appointments directly on the GCC website: <http://www.cmu.edu/gcc>. You may also visit the GCC website to find out about communication workshops offered throughout the academic year.

To find out more about any of the ways the GCC can help you, please email them at gcc-cmu@andrew.cmu.edu.

University Policy on Cheating and Plagiarism

You are expected to read and attend to the information in - [University Policy on Academic Integrity](#). The full policy is available by clicking the hyperlinked text above. Additional information about the university process for handling violations and links to resources is also available via this comprehensive website: <http://www.cmu.edu/academic-integrity/index.html>.

It is extremely important that the home-works, assignments, papers and tests that you turn in during the course reflect your own understanding. To copy answers from another person not only denies you the necessary feedback on whether or not you really understand the material, but it also compromises your integrity. In addition, those who do not succumb to cheating feel that they are “getting the short end of the stick” when they see others getting away with it. For these reasons we expect everyone to behave with integrity. It is also important that the work represents your work. Thus, any unauthorized assistance in doing the course project or homework is also considered cheating.

In this class, without explicit permission of the instructor, the following do not count as original work and would constitute cheating:

- Turning in the same or largely similar paper to another class or classes.
- Joint work with another student on a problem set or final project.
- Copying material from the web without citing it correctly.
- Plagiarism, including – copying images, graphs, and tables from published work.
- Failure to correctly cite material produced by others regardless of whether it appeared in a blog, news article, web-post, journal publication, book, etc.
- Failure to correctly cite previously published works by yourself.
- Utilizing source code developed by others or drawn from the web for your project without explicit prior permission of the instructor, and appropriate reference.

Note, papers may be assessed using automatic tools for plagiarism detection.

[Course Outline](#)

Lecture I: Introduction – What is Dynamic Network Analysis

Carley, K.M. Chapter 1, 6.1

Wasserman, S. & Faust, K. Chapters 1(1.1,1.2,1.3,1.4), 2 (2.1, 2.2, 2.3) and 3.1 and 3.2

Carley, K. M., 2004, Dynamic Network Analysis. In R. Breiger, K. M. Carley & P. Pattison (Eds.), *Dynamic Social Network Modeling and Analysis: 2002 Workshop Summary and Papers* (pp. 133-45). Washington, DC: National Academies Press.

Lecture II: Network Elite

Carley, K.M. Chapter 2

Wasserman & Faust, Chapter 5

Borgatti, S. P. (2004). The Key Player Problem. In R. Breiger, K. M. Carley & P. Pattison (Eds.), *Dynamic Social Network Modeling and Analysis: 2002 Workshop Summary and Papers* (pp. 241-52). Washington, DC: National Academies Press.

Borgatti, Stephen P., 2005. "Centrality and network flow." *Social networks* 27(1): 55-71.

Philip Bonacich, Power and Centrality: A Family of Measures, *American Journal of Sociology*, 1987, 92(5):1170-82.

Freeman, L.C. 1979. Centrality in social networks: Conceptual clarification. *Social Networks*. 1: 215-239

Lecture III: Groups

Carley, K.M. Chapter 4

Wasserman & Faust, Chapter 7,8,9,10,(12 --- *for reference only*)

Newman, M. 2004. "Detecting Community Structure in Networks." *European Physics B*:321-330.

Breiger, Ronald, Scott Boorman, and Phipps Arabie. 1975. "An Algorithm for Clustering Relational Data with Applications to Social Network Analysis and Comparison with Multidimensional Scaling." *Journal of Mathematical Psychology* 12:328-383.

Davis, George, and Kathleen Carley. 2008. "Clearing the FOG: fuzzy overlapping groups for social networks." *Social Networks* 30:201-212.

Lecture IV: Network Topology

Newman, Mark. 2003. "The Structure and Function of Complex Networks." *SIAM Review* 45:167-256.

- Borgatti, Stephen, and Martin Everett.** 1999. "Models of Core/Periphery Structures." *Social Networks* 21:375-395.
- Erdos, Paul, and Alfred Renyi.** 1959. "On Random Graphs I." *Publicationes Mathematicae Debrecen* 6:290-297.
- Kleinberg, Jon.** 1999. *The Small World Phenomenon: an algorithmic perspective.* Cornell Computer Science Department: Cornell University.
- Barabasi, Albert-Laszlo, and Eric Bonabeau.** 2000. "Scale-Free Networks." *Scientific American* 288(5):50-59.

Lecture V: Compare and Contrast Networks

Carley, K.M. Chapter 6.3

Wasserman & Faust, Chapter 15

- Krackhardt, David.** 1988. "Predicting with Networks: Nonparametric Multiple Regression Analysis of Dyadic Data." *Social Networks* 10:359-381.
- Anderson, Carolyn, Stanley Wasserman, and Bradley Crouch.** 1999. "A P* Primer: logit models for social networks." *Social Networks* 21:37-66.
- Robins, Garry, Pip Pattison, Yuval Kalish, and Dean Lusher.** 2007. "An Introduction to Exponential Random Graph (p*) Models for Social Networks." *Social Networks* 29:173-191.

Lecture VI: Issues of Analysis and Inference

- Borgatti, Stephen, Kathleen Carley, and David Krackhardt.** 2006. "On the Robustness of Centrality Measures under Conditions of Imperfect Data." *Social Networks* 28:124-136.
- Anderson, Brigham, Carter Butts, and Kathleen Carley.** 1999. "The Interaction of Size and Density with Graph-Level Indices." *Social Networks* 21:239-267.
- Banks, David, and Kathleen Carley.** 1994. "Metric Inference for Social Networks." *Journal of Classification* 11:121-149.
- Bernard, H.R., Killworth, P. & Sailer, L.** 1981. Summary of research on informant accuracy in network data, and on the reverse small world problem. *Connections* 4(2):11-25.

Lecture VII: Data Discovery, Missing Data, Sampling

- Gjoka, M., Butts, C. T., Kurant, M., & Markopoulou, A.** 2011. Multigraph sampling of online social networks. *Selected Areas in Communications, IEEE Journal on,* 29(9), 1893-1905.
- Gjoka, M., Kurant, M., Butts, C. T., & Markopoulou, A.** 2011. Practical recommendations on crawling online social networks. *Selected Areas in Communications, IEEE Journal on,* 29(9), 1872-1892.
- Airoldi, E. M., & Carley, K. M.** 2005. Sampling algorithms for pure network topologies: a study on the stability and the separability of metric embeddings. *ACM SIGKDD Explorations Newsletter,* 7(2), 13-22

Wei Wei, Kenneth Joseph, Huan Liu and Kathleen M. Carley, 2016, "Exploring Characteristics of Suspended Users and Network Stability on Twitter." *Social network analysis and mining*, 6:51.

Lecture VIII: Meta-Networks

Carley, K.M. Chapter 3

Kathleen M. Carley, 2002, "Smart Agents and Organizations of the Future" *The Handbook of New Media*. Edited by Leah Lievrouw and Sonia Livingstone, Ch. 12, pp. 206-220, Thousand Oaks, CA, Sage.

L Tang, H Liu, J Zhang, Z Nazeri, 2008, "Community evolution in dynamic multi-mode networks" *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* Pages 677-685

Lecture IX: Link Inference & From Links to Networks

Granovetter, M.S., 1973. The Strength of Weak Ties. *American Journal of Sociology* 78, 1360–1380.

Park, Han Woo, and Mike Thelwall. 2003. "Hyperlink Analyses of the World Wide Web: A Review." *Journal of Computer Mediated Communication* 8.

S. Lehman and A.D. Jackson, 2005, Live and Dead Nodes. *Computational and Mathematical Organization Theory*. 11(2): 161-170.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, 415-444.

Lecture X: Network Dynamics I (Simulation)

Carley, K.M. Chapter 7.1, 7.2

Friedkin, Noah and Eugene Johnsen. 1999. Social Influence Networks and Opinion Change. *Advances in Group Processes* 16: 1-29.

Johnson, Jeff, Palinkas, Lawrence, and Boster, James. 2004. Informal social roles and the evolution and stability of social networks. In R. Breiger, K. M. Carley & P. Pattison (Eds.), *Dynamic Social Network Modeling and Analysis: 2002 Workshop Summary and Papers* (pp. 121-32). Washington, DC: National Academies Press.

Watts, Duncan. 1999. Networks, dynamics, and the small world phenomenon. *American Journal of Sociology*, 105(2), 493-527

Lecture XI: Network Dynamics II (Inference)

Carley, K.M. Chapter 6.2,6.4,6.5

Wasserman & Faust, Chapter 17(17.2)

McCulloh, Ian, and Kathleen Carley. 2008. *Social Network Change Detection*. Technical report number CMU-CS-08-116. Carnegie Mellon University School of Computer Science: Carnegie Mellon University, Pittsburgh PA.

Snijders, Tom, Philippa Pattison, Garry Robins, and Mark Handcock. 2006. "New Specifications for Exponential Random Graph Models." *Sociological Methodology* 36:99-153.

Snijders, Tom, Christian Steglich, and Michael Schweinberger. 2007. "Modeling the Co-Evolution of Networks and Behavior." Pp. 41-72 in *Longitudinal Models in the Behavioral and Related Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lecture XII: Big Data Analytics

Lorenzo De Stefani, Alessandro Epasto, Matteo Riondato, and Eli Upfal. 2016. TRIÈST: Counting Local and Global Triangles in Fully-Dynamic Streams with Fixed Memory Size. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*. ACM, New York, NY, USA, 825-834.

Pfeffer, Jürgen & Carley, Kathleen M. 2012. k-Centralities: Local Approximations of Global Measures Based on Shortest Paths. *Proceedings of the WWW Conference 2012, 1st International Workshop on Large Scale Network Analysis (LSNA 2012)*, Lyon, France

Kas, M., Carley, K. M., & Carley, L. R. (2015). An Incremental Algorithm for Updating Betweenness Centrality and k-Betweenness Centrality and Its Performance on Realistic Dynamic Social Network Data. *Social network analysis and mining (SNAM)*. Springer Vienna. Vol. 4, No. 1, pp. 234-256, Jan 9, 2015.

Leskovec, Jure, & Faloutsos, C. (2007, June). Scalable modeling of real graphs using kronecker multiplication. In *Proceedings of the 24th international conference on Machine learning* (pp. 497-504). ACM.

Lecture XIII: Application: Social Media & BotNets

Carley, Kathleen M., Pfeffer, Jürgen, Liu, Huan & Morstatter, Fred (2014). Embassies Burning: Toward a Near Real Time Assessment of Social Media Using Geo-Temporal Dynamic Network Analytics. In: *Social Network Analysis and Mining* 4(1). 1-23.

Boshmaf, Yazan, et al. "The socialbot network: when bots socialize for fame and money." *Proceedings of the 27th Annual Computer Security Applications Conference*. ACM, 2011.

Benigni, Matthew & Carley, Kathleen, 2017, The Low Cost of Online Credibility: Socialbot Networks and Their Manipulation of Public Opinion on Twitter, Under Review

Lecture XIV-XV: Application: Organizations

Krackhardt, David, and Daniel Brass. 1994. "Interorganizational Networks: the micro side." Pp. 207-229 in *Advances in Social Network Analysis: research in the social and behavioral sciences*, S. S. Wasserman & J. Galaskiewicz (Eds.). Thousand Oaks, CA: Sage.

Burt, Ronald. 1992. "The Social Structure of Competition." Chapter 2 in *Structural Holes*. Harvard University Press, Boston MA (pp. 57-89).

Cross, R. L., & Parker, A. (2004). *The hidden power of social networks: Understanding how work really gets done in organizations*. Harvard Business Press.

Tsai, Wenpin, and Sumantra Ghoshal. 1998. "Social capital and value creation: The role of intrafirm networks." *Academy of management Journal* 41(4): 464-476.

Lecture XV: Application: TBD

Lecture XVI: Application: Counter-Terrorism

http://www.slate.com/articles/news_and_politics/searching_for_saddam/2010/02/searching_for_saddam.html

Carley, Kathleen, Ju-Sung Lee, and David Krackhardt. 2002. "Destabilizing Networks." *Connections* 24:79-92.

Enders, Walter, and Paan Jindapon. 2010, "Network externalities and the structure of terror networks." *Journal of Conflict Resolution* 54(2): 262-280.

John Horgan, Michael Kenney, Cale Horne, Peter Vining, Kathleen M. Carley, Michael Bigrigg, Mia Bloom, and Kurt Braddock, 2014, "Competitive Adaptation in Terrorist Networks: Preliminary Findings From an Islamist Case Study," in *Counter-terrorism & Hostile Intent: Human Factors Theory and Application*, edited by Alex Stedmon, Glyn Lawson & Rose Saikayasit (London: Ashgate, 2014).

Lecture XVII: Application: Health

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine*, 357(4), 370-379.

Scanfeld, D., Scanfeld, V., & Larson, E. L. (2010). Dissemination of health information through social networks: Twitter and antibiotics. *American journal of infection control*, 38(3), 182-188.

Jacqueline A. Merrill, Kathleen M. Carley, Mark G. Orr, Christie Y. Jeon and Jonathon Storricks, 2012, "Patterns of Interaction Among Local Public Health Officials and the Adoption of Recommended Practices," *Frontiers in Public Health Services and Systems Research*, 1:1, Article 6. Available from: <http://uknowledge.uky.edu/frontiersinphssr/vol1/iss1/6>

Lecture XVIII-XIX: Application: Diffusion

Watts, Duncan, and Peter Sheridan Dodds. 2007. "Influentials, Networks, and Public Opinion Formation." *Journal of Consumer Research* 34:441-458.

Kempe, David, Jon Kleinberg, and Éva Tardos. 2003. "Maximizing the spread of influence through a social network." In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 137-146. ACM.

Newman, Mark EJ. 2002. "Spread of epidemic disease on networks." *Physical review E* 66(1): 016128.

Granovetter, M., 1978. Threshold Models of Collective Behavior. *American Journal of Sociology* 83, 1420-1443.

Lecture XX-XXI: Network Text Analysis

Carley, K.M. Chapter 8

Carley, Kathleen, 1997, "Extracting Team Mental Models Through Textual Analysis." *Journal of Organizational Behavior*, 18: 533-538.

Diesner, J., Frantz, T.L., Carley, K.M., 2005. Communication Networks from the Enron Email Corpus "It's Always About the People. Enron is no Different". *Comput. Math. Organ. Theory* 11, 201-228.

Somers, M.R., 1994. The narrative constitution of identity: A relational and network approach. *Theory and society* 23, 605-649.

Smith, Marc A., Lee Rainie, Ben Shneiderman, and Itai Himelboim. 2014. "Mapping Twitter topic networks: From polarized crowds to community clusters." Pew Research Center 20.

Lecture XXII-XXVI: Final Presentations

Lecture XXVII LAST: The Future of Network Science

Carley, K.M. Chapter 9

Carley, Kathleen, and Allen Newell. 1994. "The Nature of the Social Agent." *Journal of Mathematical Sociology* 4:221-262.

Kenneth Joseph, Kathleen M. Carley, David Filonuk, Geoffrey P. Morgan, and Jürgen Pfeffer, to appear. Arab Spring: From News Data to Forecasting. *Social Networks and Mining*.

Butts, Carter, and Kathleen Carley. 2000. *Spatial Models of Large-Scale Interpersonal Networks*. Thesis Preprint. Pittsburgh, PA: Carnegie Mellon University.