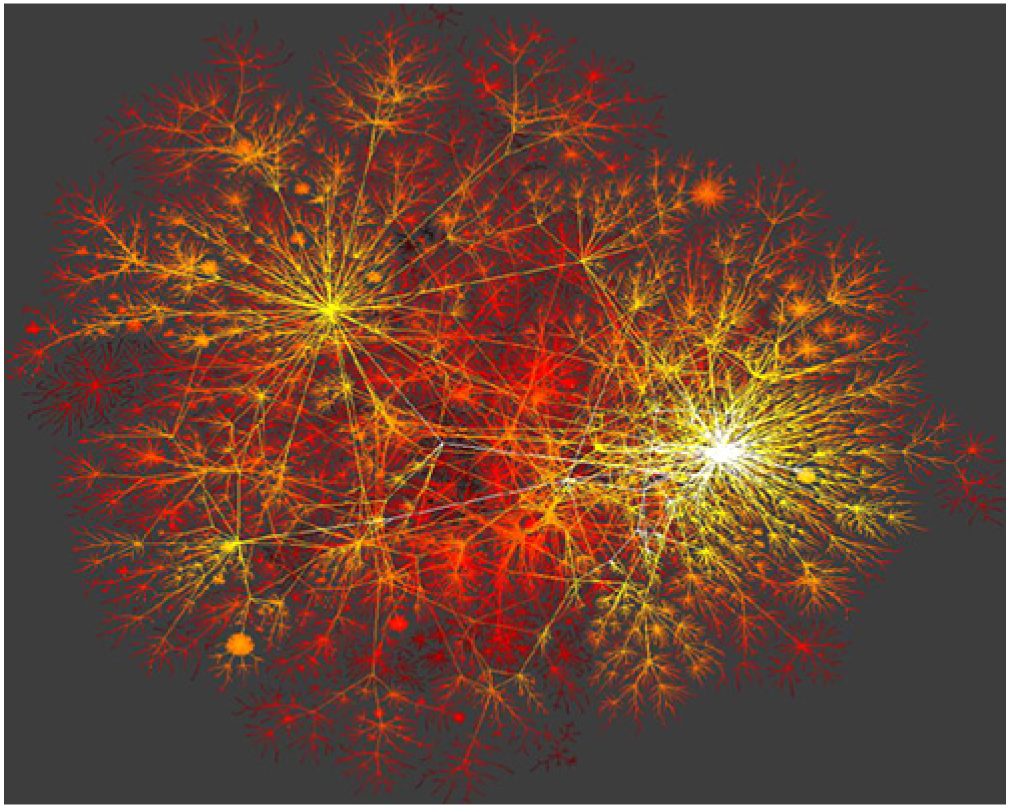
CS 7280
Network Science: Methods and Applications

Fall 2016



Instructor and classroom

Instructor: Constantine Dovrolis
Lectures: Tuesday and Thursday 1:35-2:55
Classroom: CCB 053
Office: KACB 3346

Office hours: Tuesday and Thursday, we can talk right after class (or by appointment)
Instructor's email: constantine@gatech.edu

Teaching assistant

Teaching assistant: Kamal Shadi
Office hours: TBA
TA's email: TBA

Course objectives

It is often the case that *complex systems*, both living and man-made, can be represented as static or dynamic *networks* of many interacting components. These components are typically much simpler in terms of behavior or function than the overall system, implying that the additional complexity of the latter is an *emergent network property*.

Network science is a new discipline that investigates the topology and dynamics of such complex networks, aiming to better understand the behavior, function and properties of the underlying systems.

The applications of network science cover physical, informational, biological, cognitive, and social systems. In this course, we will study algorithmic, computational, and statistical methods of network science, as well as applications in communications, biology, ecology, brain science, sociology and economics. The course will go beyond the strictly structural concepts of small-world and scale-free networks, focusing on dynamic network processes such as epidemics, synchronization, or adaptive network formation.

Course prerequisites

The course hopes to attract students from different academic backgrounds and research interests (including math, physics, engineering, biology, neuroscience or sociology). Consequently, the instructor will try to keep the course as ``self-contained'' as possible. However, some basic familiarity with calculus, probability, and programming will be necessary. Additionally, students will be free to choose course projects that are closer to their background.

References

Together with some research papers, we will cover specific chapters from the following three textbooks:

- A-L. Barabási , Network Science , available online, 2015.
- M.E.J. Newman, Networks - An introduction , Oxford Univ Press, 2010.
- D. Easley and J. Kleinberg, Networks, Crowds and Markets , Cambridge Univ Press, 2010 (also available online).

The following books will be useful references in certain parts of the course:

- R. Cohen and S. Havlin, Complex Networks - Structure, Robustness and Function , Cambridge Univ Press, 2010.
- M.O. Jackson, Social and Economic Networks , Princeton Univ Press, 2008.
- A. Barrat, M. Barthelemy and A. Vespignani, Dynamical Processes on Complex Networks, Cambridge Univ Press, 2008.
- E. Kolaczyk, Statistical analysis of network data, Springer, 2009.
- S. Wasserman, K. Faust, Social Network Analysis: Methods and Applications, Cambridge Univ Press, 1994.
- P. Van Mieghem, Graph Spectra for Complex Networks, Cambridge Univ Press, 2011.
- R. Diestel, Graph Theory (4th edition), Springer, 2010.
- R.K.Ahuja and T.L.Magnanti, Network Flows: Theory, Algorithms, and Application , Pearson, 1993.

Syllabus and course structure

The course will consist of **two types of lectures.** The first type will cover some **fundamental** concepts, metrics, algorithms, and analytical results. This material is described in several textbooks and we will not need to rely on research papers. The second type will focus on recent **applications** of network science in brain networks, biological networks, social networks, etc. In this case we will rely quite a bit on research papers that will be added (each week) at this course website.

- Course overview - Background
  - What is (not) network science?
  - History and relation to graph theory, physics, sociology, and other disciplines
  - Examples of networks from different application domains

  1. What is network science? by Ulrik Brandes et al. (Editorial from a new Network Science journal)
  2. The architecture of complexity by Albert-László Barabási (if you have time, please also read the book "Linked" by the same author)
  3. Why Model? by J.Epstein (a short but very nice discussion of why we need abstract modeling even in the era of "big data")
  4. Networks in Biology: Network Inference, Analysis, and Modeling in Systems Biology by Reka Albert
  5. Networks in Neuroscience: Complex brain networks: graph theoretical analysis of structural and functional systems by Ed Bullmore and Olaf Sporns
  6. Networks in Social Science: Network Analysis in the Social Sciences by Stephen Borgatti et al.
  7. Networks in Economics: Economic Networks: The New Challenges by Frank Schweitzer et al.
  8. Networks in Ecology: Networks in ecology by Jordi Bascompte
  9. Networks and the Web: Web science: an interdisciplinary approach to understanding the web by James Hendler et al.
  10. Networks and the Internet: Network Topologies: Inference, Modelling and Generation by Hamed Haddadi et al.

- Graph concepts and metrics

- Paths, components, degree distribution, degree correlations, clustering
- Various centrality metrics
- Metrics for signed, weighted and spatial networks
- Spectral metrics

1. Please read Chapters 6, 7, and 8 from the book "Networks - an introduction" by M.E.J.Newman (see above). If you do not have access to that book, you can read sections 1, 2, 3, and 6 from the following review by the same author: The Structure and Function of Complex Networks by M.E.J.Newman.
2. An application paper, comparing different centrality metrics in the context of spatial networks: Centrality measures in spatial networks of urban streets by P.Crucitti et al.
3. Metrics for weighted networks and an application in air travel networks: The architecture of complex weighted networks by A.Barrat et al.
4. An application paper, focusing on brain structural networks (uses k-core decomposition among other methods): Mapping the Structural Core of Human Cerebral Cortex by P.Hagmann et al.
5. Optional reading: A Faster Algorithm for Betweenness Centrality by U.Brandes (Betweenness centrality and how to compute it efficiently)

- Properties of many real networks
  - Small-world property
  - Scale-free property and heavy-tailed degree distributions
  - Hierarchy - Modularity
  - Core-periphery structure and bow-tie (or hourglass) networks
  - Network motifs

- Network models
  - Random networks
  - Watts-Strogatz model
  - Preferential attachment and its many variants
  - Kleinberg's duplication-based model
  - Optimization-based network formation models (e.g., HOT)

- Community detection methods
  - Graph partitioning
  - Modularity maximization methods
  - Hierarchical divisive and agglomerative methods
  - Overlapping communities, dynamic communities, and other variations
  - Properties of communities in real-world networks

- Statistical (or machine learning) methods in network science - network inference methods
  - Network sampling methods
  - Learning a network model from data

  - Network inference based on cross-correlations
  - Inferring causal interactions

- Dynamics of networks
  - Percolation and network resilience to random and targeted attacks
  - Growth
  - Densification
  - Rewiring

- Dynamics on networks: epidemics and other spreading phenomena
  - Network epidemics and epidemic threshold (SI, SIS, SIR models)
  - Immunization strategies
  - Identification of major spreaders
  - Computational network epidemiology

- Dynamics on networks: social influence and cascades
  - Social networks and influence/contagion
  - The Linear threshold model and the Independent cascades model
  - Empirical studies in information and behavior spreading
  - Seeding strategies on how to maximize influence

- Dynamics on networks: (mostly) search and synchronization
  - Decentralized search on networks
  - Synchronization on networks
  - Networks with capacity constraints and overload-based failures

- Dynamics on networks: controlling networks and co-evolutionary processes
  - Controling networks
  - Coevolutionary dynamics in opinion/consensus formation
  - Coevolutionary dynamics in epidemics
  - Coevolutionary effects in population dynamics

- Advanced topics (if time allows)
  - Interdependent and layered networks
  - Temporal networks
  - Games on networks
  - Network formation games

Student projects and milestones

Course projects will be of two different types: projects (projects for and PhD-student projects.

Projects for Computer Science MS students: Since this course is an elective for MS students taking the Systems Specialization, the MS-student projects will be oriented towards design and implementation. The student(s) will develop software that can perform some form of network analysis or modeling, able to handle very large networks. For example, a good course project of this type may be the implementation and application of a previously published algorithm for community detection in very large networks. MS-student projects will be evaluated based on their functionality and efficiency.

PhD-student projects: Every project of this type will be based on an individual research paper, in the broader area of network science, chosen by the student(s). The goal of the project will be to reproduce some of the qualitative results in that paper based on different data, a different method (that the student develops for this course or that is already published). In other words, the course projects will not try to solve an original research problem -- instead, the goal is to reproduce published research results but under different conditions, inputs or methods. Please note: the selected research paper cannot be something that was previously published by the student or the student's group.

Projects for students from other academic programs (e.g., MS in Analytics): These projects will need to be relevant to the student's academic program. For example, a student pursuing an MS in Analytics will need to design and perform the analysis of a network dataset from a certain perspective that the student has chosen. This exercise may not require software development but it will certainly require significant work with R, Matlab or other related data analysis tools.

Groups of two students are ideal but individual projects are also acceptable. Larger teams will need instructor approval.

All projects will be presented in class during Finals Week.

Project milestones

1. October 1: project proposal
2. November 1: progress report
3. December 1: final paper due

Grading

- *In-class quizzes* : 10%
- *Homeworks (five of them, equal-weight)*: 50%
- *Project paper and final presentation*: 40%

Network datasets

There are dozens of sites that provide pointers to network datasets. The following is just a small subset:

- "Awesome public datasets" index (very extensive -- some of them are network-related)
- KDnuggets data index (very extensive - some of them are network-related)
- NetWiki index
- ICON index
- FODAVA
- Crawdad (mostly related to wireless networks)
- Mark Newmann's pointers
- Eric Kolaczyk's pointers
- Albert Barabási 's pointers
- The Pajek pointers
- Alex Arenas' pointers
- Jure Leskovec's pointers
- Pointers to additional publicly available datasets are very welcomed

---

Network analysis and visualization tools

The following pointers provide network analysis tools that you can use in course projects:

- Gephi
- Cytoscape
- NetworkX
- Infomap
- Igraph
- Statnet
- Network Workbench
- Pajek network visualization (Windows)
- Jung network analysis
- GraphViz
- Matlab's Random Boolean Networks (RBN) toolbox
- Pointers to additional network analysis and visualization tools are very welcomed

---