NETWORK SCIENCE PHD PROGRAM HANDBOOK

ACADEMIC YEAR 2018-19

ACADEMIC DIRECTOR

David Lazer, PhD
Distinguished Professor
College of Social Sciences and Humanities
and College of Computer and Information Science

CONTACT

Mark Giannini
Graduate Program Coordinator
m.giannini@northeastern.edu
617-373-8856

Network Science Institute
177 Huntington Avenue
Boston, MA 02115

netsi | Northeastern University
Network Science Institute

CONTENTS

NETWORK SCIENCE PROGRAM FACULTY MEMBERS

ALBERT-LÁSZLÓ BARABÁSI

László Barabási is the Robert Gray Dodge Professor of Network Science and a Distinguished University Professor of Physics. He investigates biological networks, science of success, applications of control theory to networks, development of network models of resiliency in systems.

NICK BEAUCHAMP

Nick Beauchamp is an Assistant Professor in the Department of Political Science. He investigates how political opinions form and change as a result of discussion, deliberation and argument in political domains using techniques from machine learning, automated text analysis, and social network analysis.

TINA ELIASSI-RAD

Tina Eliassi-Rad is an Associate Professor of Computer Science. Her research interests include (a) statistical relational learning and graph mining, (b) axiomatic approaches to descriptive measures on graphs, and (c) unifying the physics of networks with the mining of graphs. Applications of interests include anomaly detection, cultural analytics, cyber situational awareness, and ranking systems.

DIMA KRIOUKOV

Dima Krioukov is an Associate Professor in the Departments of Physics, Mathematics, and Electrical & Computer Engineering. He develops novel applications of geometry and physics to the analysis of massive networks, including those involved in navigation/routing in the Internet, neuroscience, and cosmology.

DAVID LAZER

David Lazer is a Distinguished Professor in the Department of Political Science and College of Computer and Information Science, and co-director of the NULab for Texts, Maps, and Networks. His research interests include group learning in technology-mediated environments; consensus and opinion formation in groups, particularly in political settings, or pertaining to governance.

CHRISTOPH RIEDL

Chris Riedl is an Assistant Professor for Information Systems at the D'Amore McKim School of Business. He employs business analytics and data science to investigate research questions about group-decision making, network science, and social media, and develops novel computational approaches to study collective intelligence mechanisms.

SAMUEL SCARPINO

Samuel Scarpino is an Assistant Professor for Marine & Environmental Sciences and Physics in the College of Science. His research involves investigating questions at the intersection of network science and human behavior, whose work spans a broad range of topics, including: infectious disease modeling, forecasting in complex systems, genetic topology of disease, and decision making under uncertainty.

ALESSANDRO VESPIGNANI

Alex Vespignani is the Director of the Network Science Institute and Sternberg Family Distinguished University Professor with interdisciplinary appointments in the College of Computer and Information Science, College of Science and the Bouvé College of Health Sciences. His research interests include complex systems & networks; and the data-driven computational modeling of epidemics.

BROOKE FOUCAULT WELLES

Brooke Welles is an Associate Professor of Communication Studies in the College of Arts, Media and Design. Her research examines how social networks shape behavior, including how individuals identify resources within their social networks and leverage them to achieve personal and organizational goals.

PROGRAM OVERVIEW

Northeastern University offers a full-time doctoral program culminating in a PhD in Network Science. The objective of the program is to train a select group of students to become experts in the interdisciplinary field of network science. Network Science research covers a broad range of topics, including: Control of Networks, Computation Social Science, Biological Networks, Spreading and Influence, Group-Decision Making, Data and Graph Mining, and Network Geometry.

COURSE REQUIREMENTS

Students must take a minimum of 32 credit hours of course work. Required courses work include the following: 1) Three foundational courses in network science—Complex Networks and Applications (PHYS 5116); Network Science Data (PHYS 7331); and Dynamical Processes in Complex Systems (PHYS 7335). 2) A minimum of one of three supplemental courses in network science—Network Science Data 2 (PHYS 7332); Social Networks (POLS 7334); or Data Mining Techniques (CS 6220). 3) A minimum of 12 semester hours of elective course work. Electives are dependent on a student's area of research and subject to prior approval by their faculty advisor. 4) Two research courses (NETS 8984), for a total of four semester hours, with core faculty of the program.

The graduate program may recommend additional course work based on student research interests and in consultation with the student's faculty advisor.

NETWORK SCIENCE PHD PROGRAM - SAMPLE COURSE PLAN

YEAR	FALL	S.H.	SPRING	S.H.
1	PHYS 5116 Complex Networks and Applications	4	PHYS 7332 Network Science Data 2	4
	PHYS 7331 Network Science Data 1	4	POLS 7334 Social Network Analysis	4
	NETS 8941 Network Science Literature Review	2	CS 6220 Data Mining	4
2	PHYS 7335 Dynamical Processes in Complex Networks	4	Elective (2)	4
	Elective (1)	4	Elective (3)	4
	NETS 8984 Research	2	NETS 8984 Research	2
3	Qualifying Exam		Comprehensive Exam (Dissertation Proposal)	
4 & 5	Dissertation Continuation		Dissertation Continuation	

COURSE WAIVERS

Course waivers may be accepted toward the PhD degree course requirements, though they will not change the number of credits required for the program. The student must be able to demonstrate a graduate level of understanding of the course material they are seeking to waive. Course waivers must be approved by the Academic Director. In cases where a course waiver is granted for a required course, the student is expected to take an additional elective to meet the programs credit requirement.

ACADEMIC PROGRESS

Satisfactory progress in the program will be ongoing and formally evaluated at the end of both the first and second years of the program. Students are expected to maintain a cumulative GPA of 3.000 or better in all course work. Students are not allowed to retake courses without the express permission of the academic director. A student who does not maintain the 3.000 GPA or is not making satisfactory progress on their dissertation research may be recommended for termination by the graduate program committee.

In addition to satisfactory progress in coursework, students are expected to excel in their research obligations, under the direction and supervision of their faculty advisor. Failure to meet research obligations may result in the loss of funding, termination form the program, or both. Satisfactory performance of research obligations will be ongoing and formally evaluated at least once per year.

PHD CANDIDACY

A student is considered a PhD candidate upon completion of all required coursework with a minimum cumulative GPA of 3.0, satisfactory completion of the Qualification Exam, and satisfactory completion of the Comprehensive Exam.

QUALIFICATION EXAM

The Qualification Exam will be an oral examination of the material during the students' coursework. The exam will be an hour in length and consist of questions selected by Network Science faculty who comprise the Qualifying Examination and Dissertation Committee. Students will receive 50 – 80 potential questions, which they must be prepared to answer, one month before the exam. The exam will consist of a subset of these questions. The Qualifying Exam is will be offered twice annually, in the fall and spring term. All students are required to initially sit for the exam in the fall, typically in their third year of the PhD program. Students who do not pass the Qualifying Exam on their first attempt are expected to retake the exam in the spring term. Students may sit for the Qualifying exam no more than twice.

Students who fail to complete the Qualifying Examination but who have completed all the PhD program's required course work with a cumulative GPA of 3.000 or better will be awarded a terminal Master of Science in Network Science degree. Note that no students will be admitted directly into the network science program for receipt of a master's degree.

COMPREHENSIVE EXAM

Students must submit a written dissertation proposal to the Dissertation Committee. The proposal (with the aid and approval of their dissertation advisor) will outline a plan to carry out new and original research. The proposal should identify relevant literature, the research problem, the research plan, and the potential impact on the field. An oral presentation of the proposal will be made in an open forum before a public audience and the Dissertation Committee, followed by questions from non-committee members. The written proposal must be given to committee member at least two weeks prior to the oral presentation. After the presentation, the student will meet with the dissertation committee to address any concerns raised in either the written proposal or the presentation. The Comprehensive Exam must precede the final dissertation defense by at least one year.

DISSERTATION DEFENSE

The proposed final draft of the dissertation should be distributed to each dissertation committee member at least two weeks prior to the defense. The program will post public notification of the dissertation defense at least two weeks prior to the defense. The defense is public and consists of three parts: student presentation of the work; public questions from non-committee members; and, after dismissal of the public, questions from committee members, all of whom must be present at the defense.

The final dissertation must adhere to the policy and formatting required by Northeastern University's College of Science. Dissertations must consist of a minimum of three research chapters deemed publishable units based on novel experiments and empirical data. The results of the dissertation defense are decided by majority vote of the dissertation committee. The student's performance will result in one of three possible actions by the committee:

A) Pass;
B) Pass with revisions; or,
C) Fail.

DISSERTATION ADVISORS

The program takes a mentoring approach whereby the graduate students are researcher assistants in faculty laboratories, with opportunities to conduct research across traditional disciples utilizing network methodologies. In the laboratory, responsibility for collaboration in research gradually shifts from the faculty advisor to the student, culminating in the student's doctoral dissertation.

Each student will have one primary faculty advisor from the Network Science Doctoral Program faculty. Students are admitted into the program with a faculty advisor. In most cases, the faculty advisor is expected to become the student's dissertation advisor. Students will be expected to select their dissertation advisors by the end of the spring semester of their second year in the program.

In rare circumstances, students can elect to change their dissertation advisor during the course of their study, but this requires the approval of the new dissertation advisor, the graduate committee, and the academic director. The dissertation advisor must be a full time tenured or tenure-track member of the Northeastern University faculty. If a student's advisor leaves Northeastern, the student may continue the research direction of the dissertation with a Northeastern faculty member serving as their co-advisor. The co-advisor must first agree to act as the student's primary advisor and then be approved by the graduate committee and academic director. The student will then have two advisors: an official member of the Northeastern faculty who will be available for research and administrative matters and the ex-Northeastern advisor. If a new advisor is appointed, the ex-Northeastern faculty member may serve as an outside member of the committee.

DISSERTATION COMMITTEE

The Dissertation Committee will consist of at least 4 members: the dissertation advisor, one Network Science faculty member (in addition to the dissertation advisor), one additional tenured/tenure-track faculty member from Northeastern, and one expert in the specific topic of research (can be from outside the university).

LEARNING OUTCOMES

1) Students will demonstrate a graduate-level understanding of foundational network science concepts, including:
    a. Comprehension of the mathematics of networks, and their applications to biology, sociology, technology and other fields, and their use in the research of real complex systems in nature and man-made systems.
    b. Essential network data mining techniques from real world datasets to networks.
    c. Statistical descriptors of networks and statistical biases.
    d. Measures and metrics of networks.
    e. Network clustering techniques.
    f. Network modeling.
    g. Understanding process modeling on networks
    h. Networks visualization.

2) Familiarity with the ongoing research in the field of Network Science.

3) Students will also demonstrate a graduate-level understanding of non-network methods that enable network research, including:
    a. Computational statistics.
    b. Data acquisition and handling.
    c. Measurement and research design.

4) Graduates will attain a critical mass of understanding of some substantive domain complementary to network science, such as physics, political science, or computer science.

5) Graduates of the program should be capable of leading and performing independent, new research projects related to network sciences.

6) Students will communicate network science concepts, processes, and results effectively, both verbally and in writing.

COURSE DESCRIPTIONS

CORE COURSEWORK

PHYS 5116 - Complex Networks and Applications

Introduces network science and the set of analytical, numerical, and modeling tools used to understand complex networks emerging in nature and technology. Focuses on the empirical study of real networks, with examples coming from biology (metabolic, protein interaction networks), computer science (World Wide Web, Internet), or social systems (e-mail, friendship networks). Shows the organizing principles that govern the emergence of networks and the set of tools necessary to characterize and model them. Covers elements of graph theory, statistical physics, biology, and social science as they pertain to the understanding of complex systems.

PHYS 7331 - Network Science Data 1

Offers an overview of data mining and analysis and techniques in network science. Introduces students to network data analysis. Presents algorithms for the characterization and measurement of networks (centrality based, decomposition, community analysis, etc.) and issues in sampling and statistical biases. Reviews visualization algorithms and specific software tools. Offers students an opportunity to learn about working with real-world network datasets.

PHYS 7335 - Dynamical Processes in Complex Networks

Immerses students in the modeling of dynamical processes (contagion, diffusion, routing, consensus formation, etc.) in complex networks. Includes guest lectures from local and national experts working in process modeling on networks. Dynamical processes in complex networks provide a rationale for understanding the emerging tipping points and nonlinear properties that often underpin the most interesting characteristics of socio-technical systems. The course reviews the recent progress in modeling dynamical processes that integrates the complex features and heterogeneities of real-world systems.

NETS 8941 - Network Science Literature Review Seminar

Critically evaluates recent articles in the academic literature surrounding topics and applied research in network science. May be repeated up to three times.

NETS 8984 - Research

Offers advanced students an opportunity to work with an individual instructor on a topic related to current research. Instructor and student negotiate a written agreement as to what topic(s) are covered and what written or laboratory work forms the basis for the grade. Viewed as a lead-in to dissertation research.

SUPPLEMENTAL COURSEWORK

PHYS 7332 – Network Science Data 2

Focuses on practical exercises in real network data. Offers students an opportunity to learn how to retrieve network data from the real world, analyze network structures and properties, study dynamical processes on top of the networks, and visualize networks. The main programming language used in this course is the current industry standard. This is an interdisciplinary course.

-Or-

POLS 7334 - Social Networks

Offers an overview of the literature on social networks, with literature from political science, sociology, economics, and physics. Analyzes the underlying topology of networks and how we visualize and analyze network data. Key topics include small-world literature and the spread of information and disease

-Or-

CS 6220 - Data Mining Techniques

Covers various aspects of data mining, including classification, prediction, ensemble methods, association rules, sequence mining, and cluster analysis. The class project involves hands-on practice of mining useful knowledge from a large data set.

COMMON CONCENTRATION ELECTIVES

CS 5800 - Algorithms

Presents the mathematical techniques used for the design and analysis of computer algorithms. Focuses on algorithmic design paradigms and techniques for analyzing the correctness, time, and space complexity of algorithms. Topics may include asymptotic notation, recurrences, loop invariants, Hoare triples, sorting and searching, advanced data structures, lower bounds, hashing, greedy algorithms, dynamic programming, graph algorithms, and NP-completeness.

DSSH 6301 - Introduction to Computational Statistics

Introduces the fundamental techniques of quantitative data analysis, ranging from foundational skills—such as data description and visualization, probability, and statistics—to the workhorse of data analysis and regression, to more advanced topics—such as machine learning and networks. Emphasizes real-world data and applications using the R statistical computing language. Analyzing and understanding complex data has become an essential component of numerous fields: business and economics, health and medicine, marketing, public policy, computer science, engineering, and many more. Offers students an opportunity to finish the course ready to apply a wide variety of analytic methods to data problems, present their results to nonexperts, and progress to more advanced course work delving into the many topics introduced here.

MATH 7233 - Graph Theory

Covers fundamental concepts in graph theory. Topics include adjacency and incidence matrices, paths and connectedness, and vertex degrees and counting; trees and distance including properties of trees, distance in graphs, spanning trees, minimum spanning trees, and shortest paths; matchings and factors including matchings in bipartite graphs, Hall's matching condition, and min-max theorems; connectivity, such as vertex connectivity, edge connectivity, k-connected graphs, and Menger's theorem; network flows including maximum network flow, and integral flows; vertex colorings, such as upper bounds, Brooks, theorem, graphs with large chromatic number, and critical graphs; Eulerian circuits and Hamiltonian cycles including Euler's theorem, necessary conditions for Hamiltonian cycles, and sufficient conditions; planar graphs including embeddings and Euler's formula, characterization of planar graphs (Kuratowski's theorem); and Ramsey theory including Ramsey's theorem, Ramsey numbers, and graph Ramsey theory.

PHYS 7337 - Statistical Physics of Complex Networks

Covers applications of statistical physics to network science. Focuses on maximum-entropy ensembles of networks and on applicability of network models to real networks. Main topics include microcanonical, canonical, and grand canonical ensembles of networks, exponential random graphs, latent variable network models, graphons, random geometric graphs and other geometric network models, and statistical inference methods using these models. Covers applications of maximum-entropy geometric network models to efficient navigation in real networks, link prediction, and community structure inference.

NETS 7350 - Bayesian and Network Statistics

Introduces advanced quantitative methods including maximum likelihood, hierarchical models, sampling, and network modeling. Offers students an opportunity to learn to estimate and develop models from the probabilistic and Bayesian perspective and pursue their own research project, focusing on the methodological challenges. Reviews probability and examines maximum likelihood methods for estimating regression models with continuous and categorical dependent variables. Examines a variety of procedures for sampling from posterior distributions, including grid, quadratic, Gibbs, and Metropolis sampling. Applies these methods to hierarchical modeling and other simple probabilistic models, then takes a closer look at the statistical modeling of networks as it has been developed in the social sciences, beginning with the exponential random graph model (ERGM) and finishing with the temporal SIENA model.

NETS 7341 - Network Economics

Covers seminal works in the economics of information and networks, including Akerlof, Arrow, Spence, Stiglitz, and von Hayek. Proceeds through concepts of information, its value, and measurement; search and choice under uncertainty; signaling, screening, and how rational actors use information for private advantage; strategy-given network effects; two-sided (or multisided) network effects, organizational information processing, learning, and social networks; and other micro- and macroeconomic effects such as matching markets. Although primarily a theory course, it may be of interest to any student applying information economics and network economics in academic, commercial, or government policy contexts. Expects students to produce a major paper suitable for publication or inclusion in a thesis. Requires prior completion of graduate coursework in microeconomics and mathematics at the level of introductory calculus and statistics.

CS 7260 – Visualization for Network Science

Covers the principles of information visualization in the specific context of network science. Introduces visual encoding of data and our understanding of human vision and perception; interaction principles including filtering, pivoting, aggregation; and both quantitative and human subject evaluation techniques. Covers visualization techniques for several network types, including multivariate networks with attributes for entities and relationships, evolving and dynamic networks that change over time, heterogeneous networks with multiple types of entities, and geospatial networks. Offers students an opportunity to learn about the design of layout algorithms for node-link and matrix visualizations.