HANDBOOK

NETWORK SCIENCE PHD PROGRAM @NORTHEASTERN

network science institute

PROGRAM OVERVIEW

Working across traditional boundaries, researchers have begun to accelerate the integration of theoretical ideas and research technologies under the set of ideas and approaches that have recently been referred to as Network Science. Research on network connections among multiple types and levels of "actors" offers a potentially powerful mechanism to understand the workings of complex systems across broad areas of science, including information and technology, biological systems, health and health care, local and global political and economic processes, socio-technical infrastructures and sustainability. Adopting this view requires novel evaluations of, reconfigurations in, and innovations for standard methods of theorizing, data collection and analysis.

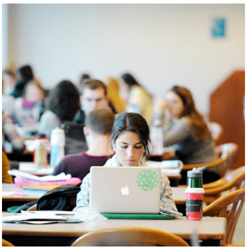
In order to provide training for the next generation of network scientists that couples deep disciplinary knowledge with inter-disciplinary Network Science, the PhD program is built on the following core principles:

> ***In-depth training in disciplines and programs is essential to interdisciplinary research. Students will be admitted in a concentration-based program. The current concentrations are focused on the physical sciences (physics); social sciences (political science); health science (epidemiology); and computer and information sciences. Additional concentrations may be added in the future.***

> ***With an aim of building an inherently interdisciplinary science and the next generation of researchers and projects. The PhD provides common, foundational training in all aspects of Network Science (e.g., approaches, languages, problems) beginning in the first year of graduate training.***

> ***The fit between theoretical/substantive questions and the use of appropriate tools and techniques for data collection and analyses represents a fundamental training element in Network Science. While we cannot expect trainees (or even faculty) to be expert, or even proficient, in all of these tools, the understanding of and respect for their potential contributions to novel interdisciplinary approaches to Network Science is paramount.***

LEARNING OBJECTIVES



Students will demonstrate a graduate-level understanding of foundational network science concepts, including:

- **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**
- **Essential network data mining techniques from real world datasets to networks**
- **Statistical descriptors of networks and statistical biases**
- **Measures and metrics of networks**
- **Network Clustering techniques**
- **Network modeling**
- **Understanding process modeling on networks**
- **Networks visualization**
- **Familiarity with the ongoing research in the field of Network Science**

Demonstrate a graduate-level understanding of non-network methods that enable Computational statistics (e.g., for social science track, a wide array of inferential methods)

Data acquisition and handling

Measurement and research design

Attain a critical mass of understanding of some substantive domain complementary to network science (such as physics, political science, computer science)

Be capable of leading and performing independent, new research projects related to network sciences

Communicate network science concepts, processes, and results effectively, both verbally and in writing

It is expected that graduates will be well-prepared to enter into a number of potential career paths including industrial research positions, government consultants, post-doctoral or junior faculty positions in academic institutions.

DEGREE REQUIREMENTS

Degree requirements are also published in Northeastern's Graduate Catalog.

Coursework is dependent on a student's area of concentration and subject to prior approval by their faculty advisor. Required coursework will include the following: Three foundational courses in Network Science (Complex Networks and Applications, Network Science Data I, and Dynamical Processes in Complex Networks); one of two approved courses (Social Network Analysis or Data Mining Techniques); twelve semester hours of elective coursework defined by their specific track; and two research courses with core faculty of the program. A minimum of 32 credit hours of coursework is required, though the graduate program committee may recommend additional coursework based on student research interests.

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 will be expected to maintain a cumulative GPA of 3.0 or better in all coursework. Students will not be allowed to retake courses. A student who does not maintain the 3.0 GPA, or is not making satisfactory progress on their dissertation research, may be recommended for termination by the Graduate Program Committee. Each student will have one primary research advisor from the Network Science Doctoral Program faculty. Students will be expected to select their research advisor by the end of the spring semester of their second year in the program.

The Dissertation Committee will consist of at least 4 members: the dissertation advisor, one additional Network Science Doctoral program faculty member, one member expert in the specific topic of research (can be from outside the university), and one additional tenured/tenure-track faculty member from the concentration department/conferring College. The dissertation advisor must be a full time tenured or tenure-track member of the Northeastern University faculty. The dissertation committee must be approved by the Graduate Program Committee and constituted no later than the end of the spring semester of the second year of the program. Students may repeat the Comprehensive Examination once if they are unsuccessful.

DEGREE CANDIDACY

A student is considered a Ph.D. degree candidate upon completion of all required coursework with a minimum cumulative GPA of 3.0, satisfactory completion of the Qualifying Examination, and satisfactory completion of the Comprehensive Examination.

QUALIFYING EXAMINATION

The Qualification Examination 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 EXAMINATION

Students must submit a written dissertation proposal to the Qualifying Examination and Dissertation Committee. The proposal should identify relevant literature, the research problem, the research plan, and the potential impact on the field. A presentation of the proposal will be made in an open forum, and the student must successfully defend it before the Qualifying Examination and Dissertation Committee. The Comprehensive Exam must precede the final dissertation defense by at least one year.

DISSERTATION DEFENSE

A Ph.D. student must complete and defend a dissertation that involves original research in Network Science. The dissertation defense must adhere to the College of Science policies, as outlined in the Northeastern University Graduate Catalog.

SAMPLE COURSE OUTLINE | NETWORK SCIENCE PHD

YR.		COURSE	S.H.
1	FALL	PHYS 5116 | Complex Networks and Applications	4
		PHYS 7331 | Network Science Data	4
	SPRING	PHYS 7331 | Network Science Data II	4
		Concentration Elective [2]	3-4
2	FALL	PHYS 7335 | Dynamical Processes in Complex Networks	4
		Concentration Elective [3]	3-4
		NETS 8984 | Research	2
	SPRING	POLS 7334 | Social Network Analysis or CS 6220 | Data Mining	4
		Concentration Elective [4, if needed]	3-4
		NETS 8984 | Research	2
		NETS 9000 | PhD Candidacy Achieved	
3	FALL	NETS 9990 | Dissertation	
	SPRING	NETS 9990 | Dissertation	
4	FALL	NETS 9996 | Dissertation Continuation	
	SPRING	NETS 9996 | Dissertation Continuation	
5	FALL	NETS 9996 | Dissertation Continuation	
	SPRING	NETS 9996 | Dissertation Continuation	

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 I

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 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.

REQUIRED ELECTIVE

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

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.

SAMPLE CONCENTRATION ELECTIVES

PHYS 7332 | Network Science Data II

Interdisciplinary course focusing on the practical exercises in real network data. Topics include how to retrieve network data from the real world, analyze network structures and properties, dynamical processes on top of networks, and visualization of networks.

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 covered 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.

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 non-
experts, and progress to more advanced course work delving into the many topics introduced here.

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.

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.

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.

ADMISSION CRITERIA AND PROCESS

Application materials include:

- **transcript(s)**
- **personal statement**
- **three letters of reference**
- **general GRE**
- **TOEFL (in the case of international applicants)**

The requirements for admission to and completion of the program conform to the University Graduate Council By-laws. Offers of admissions will be made based on the applicant's qualifications and space within the program.

Students will be accepted with a bachelor's or higher degree in any field and should have either academic or work experience demonstrating a commitment to working in network science Successful applicants to the program from non-technical disciplines will be expected to present a strong background in mathematics.

Successful applicants will be expected to have a minimum GPA of 3.30; either Verbal or Quantitative GRE General scores at 75th percentile or higher with remaining scores at 50th percentile or higher; and a minimum TOEFL of 100.

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 Professor in the Department of Political Science and College of Computer and Information Science. 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. Broadly, Scarpino is a complex systems scientist 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 Assistant 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.

Photo credits: Brooks Canaday (10), Conor Doherty (6), Mary Knox Merrill (2, 14), Matthew Modoono (4, 15) | Northeastern University

Network Science Institute
177 Huntington Ave
Boston, MA 02115

CONTACT

Mark Giannini
Program Coordinator
m.giannini@neu. edu
617 373 8856

MAILING ADDRESS

Northeastern University
360 Huntington Avenue
1010-177
Boston, MA 02115

nunetsi.org/phd

@NUnetsi

netsi

Northeastern University
Network Science Institute