COURSE: NETWORK DATA SCIENCE: MAPS, MODELS, AND ANALYSIS
DEPARTMENT: TDS – TECHNOLOGY AND DATA SCIENCE (TECNOLOGIA E CIÊNCIA DE DADOS)
PROGRAM: CMCD AE
CLASS-HOURS: ☒30 hours or ☐15 hours
PROFESSOR: Dr. Phokion 'Ion' Georgiou
SEMESTER/YEAR: 2º 2019
LANGUAGE: ENGLISH

COURSE DESCRIPTION

One of the major reasons why organizations – such as institutions, industries, markets, blocs, or simply groups of people – emerge as complex systems is because, behind the perceived complexity, there is a network that defines the interactions between the constituent entities. Uncovering the network, mapping it, tackling its complexity through models, and analyzing the results, yields greater understanding of the dynamics that render organizations complex, and opens means for improving their configuration. Since the dawn of the 21st Century, network science has emerged as one of the most productive analytical concepts that complements statistical and qualitative approaches.[1] Its effectiveness stems from having reconceptualized the focus of analysis: from entities, to the relations that bind them. The relation has become the variable of interest, and network science has developed a host of methods through which relations can be mapped, modeled, quantified, and analyzed. For these reasons, organizational network analysis has been adopted as one of the management toolkits by international consulting firms such as McKinsey, Deloitte, ATKearney, RobCross Consulting, and Maven7.[2]

Network science is the fruit of at least four forces that arose in the 20th Century: advances in graph theory,[3] social network research,[4] communications networks (including the development of the Internet)[5], and the holistic approach of ecology[6]. Currently, it is broadly divided into two schools. On the one hand, natural scientists apply network concepts and techniques to the fields of physics, chemistry, biology and so forth[7]. On the other hand, organizational theorists are interested in the applications of network science to human contexts. No matter the breadth and depth of interest in such contexts, be it global, regional, industrial, or corporate, organizational networks are formed by, and composed of, people. As such, organizational network approaches take their lead from the field of social network analysis. This is evidenced in organizational network studies of industries[8], markets[9], economics[10], politics[11], governance[12], geopolitics[13], human resource management[14], social capital[15], competition[16], the diffusion of innovations[17], and even start-ups[18]. All such applications are founded on scholarship that tackles social networks as a general category[19].

The Course draws from the methodological approaches in social network analysis applicable to organizational issues, in order to furnish students with the latest tools for doing organizational network research. Familiarization, even fluency, with specialized network software is necessary and, in this respect, the Course introduces organizational network models and analysis using the award-winning software *Pajek*. Overall, the Course provides a foundational understanding of network

dynamics; advanced abilities in modeling and analyzing network situations; a systemic perspective of the field; and, enhances decision-making skills for situations characterized by complexity and interconnectedness.

LEARNING GOALS

The course learning goals are presented in the table below, showing how they contribute to the learning goals related to the objectives of CMCDAE.

CMCDAE Objectives	Course learning goals	Level of contribution
Qualitative research methods		○○○
Quantitative research methods	The focus of this course is "network science," that is, the quantitative methods for exploring and analyzing relations. As such, it contributes decisively to the analytical equipment of quantitative methods in general. The course offers a solid foundation in network science as a quantitative research method that can effectively tackle relational quantitative data. Students learn to design, build, and analyze quantitative network models for a wide range of organizational situations, and to convert raw network data into macro and micro metrics and indices in order to measure a wide range of network issues.	●●●
Knowledge of research themes and theory	The course introduces the methodological principles of design and analysis that apply across networks in otherwise varying contexts. It also demonstrates how complexity emerges from networks, that is, how behind each complex system there is an intricate network that encodes the interactions between the system's components. The course thus contributes to an understanding of complexity science. The scope and breadth of research themes in the fields of network and complexity science will be introduced throughout the course.	●●○

CMCDAE Objectives	Course learning goals	Level of contribution
Research procedures	In terms of research procedures, the course is innovative in introducing the specificities required of network research, from data collection to analysis to presentation. It trains students in using specialized network software necessary for dealing with the complexity of network models and data. The nature of networks, as well as the state of the art in network science, allows for sophisticated exploratory research. Inferential research is somewhat more limited given that the "sample-population" dynamic evident in contexts amenable to traditional statistical treatments is not applicable to network contexts. However, the course demonstrates a number of inferential procedures that are applicable in network contexts, and therefore shows how statistics and network science complement each other.	● ● ○
Relevance and innovation in research	Network Science is one of the fastest growing exploratory and analytical approaches to research in organizations, and underpins all of complexity research. By attending this course, students effectively amplify their range of research capabilities.	● ● ●
Development of academic papers	By the end of this course, the student will be intellectually equipped to study the network science literature that addresses organizational and social networks. Publishable work will be possible only by studying complementary, and increasingly specialized, literature with the assistance of academic supervision.	● ○ ○
Other course learning goals: In general, the student will learn that network analytics, and network "big data," are based on methodology and paradigmatic principles that are necessarily different from those encountered in other contexts or other courses on quantitative analysis. The studies offered by this course equip the student to begin to undertake advanced analyses in complexity and systemic problems.		

The full description of the CMCDAE objectives, and other related information, may be found at https://rebrand.ly/cmae-eaesp (masters) e https://rebrand.ly/cdae-eaesp (doctorate).

PREVIOUS KNOWLEDGE REQUIRED

The course does not require prior knowledge. Students may, optionally, consult texts to gain some perspective on the idea of "networks."

For a general conceptual introduction to the subject matter covered in this course, students may consult any one of the following texts:

- Caldarelli G, Catanzaro M (2012) *Networks: A Very Short Introduction*. Oxford University Press: Oxford

- Kadushin C.(2011) *Understanding Social Networks: Theories, Concepts, and Findings*. Oxford University Press: Oxford

For students interested in a general analytical introduction, based on contexts that are easy to visualize, the following texts are recommended:

- Aldous JM, Wilson RJ (2000) *Graphs and Applications: An Introductory Approach*. Springer: London
- Hage P, Harary F (1984) *Structural Models in Anthropology*. Cambridge University Press: Cambridge

For an introduction to the underlying mathematical theory of networks, the following two classic texts continue to be the prime reference points:

- Harary F (1969) *Graph Theory*. Addison-Wesley: Reading, Mass.
- Harary F, Norman RZ, Cartwright D (1965) *Structural Models: An Introduction to the Theory of Directed Graphs*. Wiley: New York

An extended list of over 100 relevant complementary texts, including texts specializing in particular applications (such as economics, politics, markets, etc.), will be made available at the beginning of the course. All texts are available in the FGV library.

CONTENT/METHODOLOGY

The Course focuses on structures of human groups, such as communities, institutions, industries, markets, countries, and blocs. These structures are conceptualized as networks of ties. The ties are the main variable of interest, because they transmit behavior, services, information or materials. The patterns of ties in any one network also provide insights into the entities linked by them. Therefore, the course addresses the concept of network, introducing several types of networks and the ways in which they can be modeled and analyzed visually and computationally. Techniques that combine relational data (such as links) with nonrelational attributes of entities (such as entity-specific economic indices) are discussed. The nonrelational attributes enhance interpretations of network structure and also enable increasingly focused analyses of subnetworks (say, the trade flows between a particular set of countries situated within a wider trade bloc).

Collective norms, shared strategies, industrial cohesion, market attitudes, and similar behavior emerge from relations between organizational entities. A major concern of organizational network analysis is to investigate who is related and who is not, and why. The Course introduces a variety of techniques to detect cohesive subgroups based on the underlying structure of respective networks. These techniques comprise means to various ends, rather than an end in themselves, and are used throughout the Course as complements to advanced analyses. Furthermore, they enable tests of whether structurally delineated groups differ with respect to various nonrelational attributes. The Course tackles the differences between, and the implications of, grouping entities according to structural properties and non-network attributes. It introduces a variety of network and subnetwork measurements of cohesion, as well as analyses that help identify components according to various

criteria. Methods are also introduced for analyzing and optimizing the composition of teams, alliances, and coalitions.

Networks are structures that allow for the transport and exchange of information, services, and materials. In this perspective, familiarization with network structure helps explain diffusion of anything, from a product innovation to a disease. Some sections of networks permit rapid diffusion, whilst others act as bottlenecks. In addition, the position of specific entities in networks gives them social capital, competitive advantages, or allows them to assume a variety of brokerage roles. Such positions may put pressure on certain entities, but can also yield power and profit. The Course introduces various indices of centrality and centralization, as well as various interpretations of these important concepts along with their respective computational techniques. The distinction between an ego-centered and a socio-centered approach to centrality analyses is discussed in depth with associated modeling methodology and analytical tools. The quantitative and qualitative value of links between entities is then introduced as a means for understanding and computing various indices of social capital. A structural approach to the analysis of competition is introduced as a significant complement to other approaches toward this issue. Finally, the Course discusses and models diffusion processes that underlie social, organizational, communicative, administrative, and marketing behavior. The modeling of diffusion through networks is introduced, with a focus on investigating structural positions of entities, their relations, and their diffusion and adoption behavior.

The Course follows the pedagogic approach known as *problem-based learning*.[20] The Course content focuses on practice which enables the simultaneous emergence of theory and methods. Students are given regular practical assignments covering a range of issues examined in the course, as well as additional challenges that require some independent research. All assignments require the design of maps, models and analyses in the specialized network software *Pajek*. As such, students learn by doing: they acquire knowledge of network concepts by applying network analysis.

ASSESSMENT

40% Mid-Term Exam (assignments 1-4)
40% Final Exam (assignments 5-8)
20% Participation

The assignments are of varied difficulty and are designed to test various aspects of the course content, at times combining various analytical procedures to demonstrate the nature of advanced analyses.

Participation constitutes 20% of the final mark. The course content and the assignments are designed to elicit active student involvement throughout the course. Students are expected to maintain a regular regime of studious reading from the textbook, and other assigned material, and to demonstrate such study through their participation.

Pass mark for this course is ≥ 6.0

BIBLIOGRAPHY

There is no shortage of books that can be used as a textbook for this Course, ranging from the most basic[21] to the most advanced[22], by way of the seemingly intermediate[23]. Given the complex nature of networks, however, any such study must be complemented by training in the use of specialized analytical software. As with books, there are numerous software packages available[24]. Choosing between them should account for the extent to which the user can evaluate the procedures and analyses of a software package, thus enabling control and ownership of its results.[25] What is required is a source that offers a combination of sufficient theoretical and methodological knowledge with hands-on training in a transparent software package.

Such a source is available in its third, revised and updated edition:

COURSE TEXTBOOK: de Nooy W, Mrvar A, Batagelj V (2018) *Exploratory Social Network Analysis with Pajek*. Revised and Expanded Edition for Updated Software. 3rd Edition. Cambridge University Press: Cambridge

Throughout the course, the textbook will be referred by its acronym: ***ESNAP***.

Students must acquire the course textbook since it is central to the content and approach of this course.

The textbook covers issues applicable to the widest possible variety of contexts, but especially focuses on human contexts of interest in organizational network analysis. As for *Pajek,* this refers to an award-winning, freely available[26] software package. This software is designed specifically as a network calculator that can handle billions of vertices, and their relations, irrespective of context. It is, therefore, useful for both, abstract and empirical analyses. *Pajek*'s transparency stems from requiring active user engagement, affording precise operational oversight with consequent demystification of the black box. Furthermore, *Pajek* has a long history of published algorithms which are open to evaluation.[27] All this enables users to maintain control of their use of the software instead of being controlled by it. In addition, the software provides outstanding graphics of networks, with multiple means for manipulating their aesthetic presentation, thus allowing for sophisticated visual appreciation to complement analytical results.

COURSE SCHEDULE

All sessions are held in a computer laboratory.

Note: In the schedule, "ESNAP" refers to the course textbook (see section entitled "Bibliography").

Session	Date	Topic	ESNAP Chapter
1	11/10/2019	FUNDAMENTALS: Introduction to network modeling; Multiple relation networks; Basic automated outputs; Visualization techniques; Attributes, structural properties, and partitioning; Local, global, and contextual subnetworks; Modeling continuous properties (vectors).	1 & 2
2	18/10/2019	COHESION: Density and degree; Introduction to basic network components and cohesive subgroups; Cores and cliques.	3
3	25/10/2019	COHESION: Alliances and teams; Structual balance and clusterability.	4
4	01/11/2019	BROKERAGE: Distance; Centrality and centralization measurements, criteria, and analysis.	6
5	08/11/2019	BROKERAGE: Bridges and bi-components; ego-networks; social capital; Affiliation and brokerage roles.	7
6	22/11/2019	BROKERAGE: Introduction to diffusion modeling; Diffusion as contagion; Exposure and adoption thresholds; Critical mass and threshold lags.	8
Additional topic (subject to time availability): RANKING: Introduction to prestige; Popularity and indegree; Domains; Proximity prestige			9

It is strongly recommended that, prior to attending each session, students should have studied the associated textbook chapter and attempted the practical examples and exercises therein. Such prior preparation significantly enhances the learning process.

PROFESSOR MINI CV

This course is taught by Professor Dr Ion Georgiou, a full professor at the Fundação Getulio Vargas (Escola de Administração de Empresas de São Paulo). Among his interdisciplinary interests, he specializes in network science, graph theory, and problem structuring methods. He has lived and worked in four continents, and is proficient in five languages. He is the author of *Thinking Through Systems Thinking*, the book that sets out a complete epistemology for systemic approaches to problematic situations, along with their consequent practical and ethical implications[28]. He regularly publishes papers in top international scientific journals, three of which are of especial relevance to this course: (i) the use of network science to solve one of the 20th Century's outstanding analytical problems in decision-making[29]; (ii) the use of network science in mapping a fundamental issue at the heart of public administration[30]; and, (iii) the use of network science in uncovering the structure and consequences of decision-making in infrastructure development[31]. He is currently researching the application of network science to the composition of teams, alliances, and coalitions.

Academic Background

PhD Lancaster University, UK (Organizational Behavior)
MSc London School of Economics, UK (Operational Research)

Contacts:
E-mails: phokion.georgiou@fgv.br ; iongeorgiou@gmail.com
LinkedIn: https://www.linkedin.com/in/ion-georgiou/

REFERENCES

[1] On how network science complements other fields in general, see: Barabási A-L (2016) *Network Science*. Cambridge University Press: Cambridge.

On how it complements statistical and other mathematical approaches, see: Kolaczyk ED (2009) *Statistical Analysis of Network Data: Methods and Models*. Springer: New York; Kolaczyk ED, Csárdi G (2014) *Statistical Analysis of Network Data with R*. Springer: NY; Pattison P (1993) *Algebraic Models for Social Networks*. Cambridge University Press: Cambridge.

On how it complements qualitative research, see: Dominguez S, Hollstein B (2014) *Mixed Methods Social Networks Research: Design and Applications*. Cambridge University Press: Cambridge

[2] See, for example:

McKinsey: https://www.mckinsey.com/business-functions/organization/our-insights/the-role-of-networks-in-organizational-change

Deloitte: https://www2.deloitte.com/us/en/pages/human-capital/articles/organizational-network-analysis.html

ATKearney: https://www.atkearney.com/leadership-change-organization/network-navigation

RobCross Consulting: https://www.robcross.org/research/what-is-ona/

Maven7: http://maven7.com/

[3] Harary F, Norman RZ, Cartwright D (1965) *Structural Models: An Introduction to the Theory of Directed Graphs*. Wiley: New York; Harary F (1969) *Graph Theory*. Addison-Wesley: Reading.

[4] Moreno JL (1934) *Who Shall Survive? A New Approach to the Problem of Human Interrelations*. Nervous and Mental Disease Publishing Company: Washington DC

[5] Hafner K, Lyon M (1996) *Where Wizards Stay Up Late: The Origins Of The Internet*. Simon & Schuster: New York

[6] Worster D (1994) *Nature's Economy: A History of Ecological Ideas*. Cambridge University Press: Cambridge

[7] See, for example: Newman MEJ, Barábasi AL (2006) *The Structure and Dynamics of Networks*. Princeton University Press: Princeton

[8] Herrigel G (2000) *Industrial Constructions: The Sources of German Industrial Power*. Cambridge University Press: Cambridge; Mizruchi MS (1992) *The Structure of Corporate Political Action: Interfirm Relations and Their Consequences*. Harvard University Press: Cambridge; Mizruchi MS, Schwartz M (1987) *Intercorporate Relations: The Structural Analysis of Business*. Cambridge University Press: Cambridge; Valente, TW (2010) *Social Networks and Health: Models, Methods, and Applications*. Oxford University Press: Oxford; Mizruchi MS (1982) *The American Corporate Network 1904-1974*. Sage: New York; Heemskerk EM (2007) *Decline of the Corporate Community: Network Dynamics of the Dutch Business Elite.* Amsterdam University Press: Amsterdam; Freeland RF (2005) *The Struggle for Control of the Modern Corporation: Organizational Change at General Motors, 1924–1970*. Cambridge University Press: Cambridge; Carolan BV (2013) *Social Network Analysis and Education: Theory, Methods & Applications*. Sage: London

[9] Easley D, Kleinberg J (2010) *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press: Cambridge

[10] Jackson MO (2008) *Social and Economic Networks*. Princeton University Press: Princeton; Knoke D (2012) *Economic Networks*. Polity Press: Cambridge; Goyal S (2007) *Connections: An Introduction to the Economics of Networks*. Princeton University Press: Princeton

[11] Knoke D (1990) *Political Networks: The Structural Perspective*. Cambridge University Press: Cambridge

[12] Kogut B (2012) *The Small Worlds of Corporate Governance*. MIT Press: Cambridge

[13] Maoz Z (2010) *Networks of Nations: The Evolution, Structure, and Impact of International Networks, 1816-2001*. Cambridge University Press: Cambridge

[14] Kilduff M, Krackhardt D (2008) *Interpersonal Networks in Organizations: Cognition, Personality, Dynamics, and Culture*. Cambridge University Press: Cambridge; Friedkin NE, Johnsen EC (2014) *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*. Cambridge University Press: Cambridge; Perry BL, Pescosolido BA, Borgatti SP (2018) *Egocentric Network Analysis: Foundations, Methods, and Models*. Cambridge University Press: Cambridge; Crossley N, Bellotti E, Edwards G, Everett MG, Koskinen J, Tranmer M (2015) *Social Network Analysis for Ego-Nets*. SAGE: London

[15] Burt RS (2005) *Brokerage and Closure: An Introduction to Social Capital*. Oxford University Press: Oxford

[16] Burt RS (1992) *Structural Holes: The Social Structure of Competition*. Harvard University Press: Cambridge; Burt RS (2010) *Neighbor Networks: Competitive Advantage Local and Personal*. Oxford University Press: Oxford

[17] Moon FC (2014) *Social Networks in the History of Innovation and Invention*. Springer: New York; Rogers EM (1995) *Diffusion of Innovations*. Fourth edition. Free Press: New York

[18] Tsvetovat M, Kouznetsov A (2011) *Social Network Analysis for Startups: Finding connections on the social web*. O'Reilly Media: Sebastopol CA

[19] McCulloh I, Armstrong H, Johnson A (2013) *Social Network Analysis with Applications*. Wiley: Chichester; Wasserman S, Faust K (1994) *Social Network Analysis: Methods and Applications*. Cambridge University Press: Cambridge; Borgatti SP, Everett MG, Johnson JC (2013) *Analyzing Social Networks*. Sage: London; Kadushin C.(2011) *Understanding Social Networks: Theories, Concepts, and Findings*. Oxford University Press: Oxford; Aggarwal CC (2011) *Social Network Data Analytics*. Springer: New York; Bellotti E (2015) *Qualitative Networks: Mixed methods in sociological research*. Routledge: NY; Carrington PJ, Scott J, Wasserman S (eds.) (2005) *Models and Methods in Social Network Analysis*. Cambridge University Press: Cambridge; Hennig M, Brandes U, Pfeffer J, Mergel I (2013) *Studying Social Networks: A Guide to Empirical Research*. Campus Verlag: New York; Knoke D (2008) *Social Network Analysis*. Second Edition. Sage: London; Marschall N (2007) *Methodological Pitfalls in Social Network Analysis*. VDM Verlag: Saarbrücken; Robins GL (2015) *Doing Social Network Research: Network-based Research Design for Social Scientists*. Sage: New York; Scott JP, Carrington PJ (2011) *The SAGE Handbook of Social Network Analysis*. Sage: CA; Wasserman S, Galaskiewicz J (1994) *Advances in Social Network Analysis: Research in the Social and Behavioral Sciences*. Sage: London.

[20] Barell JF (2007) *Problem-Based Learning: An Inquiry Approach*. Second edition. Corwin Press: Thousand Oaks

[21] Prell C (2011) *Social Network Analysis: History, Theory and Methodology*. Sage: CA

[22] Estrada E (2011) *The Structure of Complex Networks: Theory and Applications*. Oxford University Press: Oxford

[23] Newman MEJ (2010) *Networks: An Introduction*. Oxford University Press: Oxford

[24] Huisman M, Van Duijn MA (2005) Software for social network analysis. In Carrington PJ, Scott J, Wasserman S (Eds.) (2005) *Models and Methods in Social Network Analysis*. Cambridge University Press: Cambridge, pp. 270-316; Huisman M, van Duijn MA (2011) A reader's guide to SNA software. In Scott J, Carrington PJ (Eds.) *The SAGE Handbook of Social Network Analysis*. SAGE: London, pp. 578-600; Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F (2011) Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology* **62**(7): 1382–1402;

[25] Ackoff R (1967) Management misinformation systems. *Management Science* **14**(4): B147-B156

[26] http://mrvar.fdv.uni-lj.si/pajek/

[27] Batagelj V (1991) Some mathematics of network analysis. Network seminar, Department of Sociology, University of Pittsburgh, January; Batagelj V (2003) Efficient algorithms for citation network analysis. University of Ljubljana, Institute of Mathematics, Physics and Mechanics, Department of Theoretical Computer Science, Preprint Series, 41, 1-27 (arXiv:cs/0309023v1, 14 September); Batagelj V, Mrvar A (1998) Pajek: a program for large network analysis. *Connections* **21**(2): 47-57; Batagelj V, Mrvar A (2008) Analysis of kinship relations with Pajek. *Social Science Computer*

Review **26**(2): 224-246; Batagelj V, Mrvar A (2014) Pajek. In Alhajj R, Rokne J (Eds.) *Encyclopedia of Social Network Analysis and Mining* New York: Springer, pp. 1245-1256; Batagelj V, Zaversnik M (2011) Fast algorithms for determining (generalized) core groups in social networks. *Advances in Data Analysis and Classification* **59**(2): 129-145; Batagelj V, Doreian P, Ferligoj A, Kejzar N (2014) *Understanding Large Termporal Networks and Spatial Networks: Exploration, Pattern Searching, Visualization and Network Evolution*. Chichester: Wiley; Batagelj V, Cerinsek M (2013) On bibliographic networks. *Scientometrics* **96**(3): 845-864.

[28] Georgiou I (2007) *Thinking Through Systems Thinking.* Routledge: London.

[29] Georgiou I, Heck J, Mrvar A (2019) The analysis of interconnected decision areas: A computational approach to finding all feasible solutions. *Group Decision and Negotiation*. (Currently awaiting print publication, and available as Online First at: https://doi.org/10.1007/s10726-018-9607-5).

[30] Georgiou I (2014) Seeing the forest for the trees: an atlas of the politics-administration dichotomy. *Public Administration Review* **74**(2): 156-175.

[31] Georgiou I (2009) Mapping railway development prospects in Brazil. *Transport Reviews* **29**(6): 685-714.