# Leading Edge Learning in Network Science



**Ralucca Gera** 

## 1 Introduction

The current work describes the method that has been used since 2014 at a the Naval Postgraduate School (NPS) in teaching MA 4404, the Structure and Analysis of Complex Networks course, primarily to USA and international officers during their master's and doctoral program. This course is taught in the Applied Mathematics Department, as part of the Network Science Academic Certificate that students can receive along with their master's or doctoral degrees in any math-related curriculum. The students interested in the course have a technical background, generally in mathematics, computer science, operations research, or engineering. Additional information on this course and how it fits within the certificate can be found at http://faculty.nps.edu/rgera/NetSci/Certificate/dist/index.html.

Researchers have taught and used network analysis since the eighteenth century, starting with the classic Seven Bridges of Königsberg problem in graph theory [31]. In many mathematics courses, students in the same curricula with the same background and interest were exposed to information building on the same mathematical prerequisites. In recent years, students with mixed backgrounds want to learn about networks, which initiated the desire to modify standard teaching, from motivation to solutions and their interpretations, in particular, taking on a guided discovery approach, asking students to experiment with networks, and discover the reasons behind what they observed, in order to support student learning (which sometimes is hard to optimally use) [18, 27, 41].

As I promote active learning in teaching, defined by Bonewell et al. [17], the challenge is to undertake learning activities that will engage students and motivate their learning, regardless of the individual background. This stimulates interest in the class, improves the learning experience, and increases their chance of successfully

R. Gera (🖂)

Naval Postgraduate School, Monterey, CA, USA

<sup>©</sup> Springer International Publishing AG, part of Springer Nature 2018

C. B. Cramer et al. (eds.), *Network Science In Education*, https://doi.org/10.1007/978-3-319-77237-0\_2

recalling and using learned ideas in the future [13, 33]. As student interest in a course increases if the course is relevant to the student, I make it pertinent by allowing students to pick individual networks that they will analyze for the rest of the course. As I teach new concepts, they apply them to this chosen network, making it both interesting to them and to the rest of the class as they see a variety of networks analyzed. As the analysis of these networks is discussed in the class, it brings the contrast and diversity needed for learning.

Since I had the freedom of designing the first course in network science at NPS, I made the following choices, which may change. The analysis tools chosen for the course incorporate choices for each type of student background: Gephi [1], Python [2], and R [3], commonly used in network science [10, 20, 24, 39, 45]. Most students use Gephi for visualization and either Python or R for analysis and constructions. A Research Project (Sect. 5.3) is incorporated in the course, as I believe that project-based learning is a good approach to education designed to engage students [15]. Moreover, through the design of the Research Project (Sect. 5.3), students learn  $L^{A}T_{E}X$ , most of them typing their theses in  $L^{A}T_{E}X$  in the quarters following the alternative to standard mathematical homework assignments: the Network Profile Summary (Sect. 5.4) in which students apply weekly learned concepts to analyze a personal network, rather than all students analyzing the same network/ data. I believe it promotes self-determination motivates student learning [28, 52].

The course generally starts by showing existing networks, as well as how big data can be modeled by networks. Currently this is of interest to most of my officers in trying to understand emerging phenomena in technology and society. Some examples of networks presented in class are online social networks, the Internet, the World Wide Web, neural networks, food webs, metabolic networks, power grids, airline networks, national highway networks, the brain, and others. These examples are complemented by networks that students choose to create based on their interests and previous experience, including: terrorist networks, the US Tesla network, the global transportation network, snapshots from YouTube, and Twitter data.

The course then proceeds with the basic generative models (random graphs, small-worlds, preferential attachment) and newer ones based on the interest of the class. I start with the Erdős–Rényi random networks that combine the just-learned concepts of graph theory with probability theory, followed by more sophisticated models of network formation, including: Milgram's 1967 experiment [53] and Watts–Strogatz small-world networks [55], the Barabási–Albert preferential attachment growing model [9] and its variants, the Molloy-Reed configuration model [47], the random geometric model [29], and other ones that are relevant at that time. In the last couple of years, the Research Project (Sect. 5.3) has built on this overview of generative models by asking students to create new ones.

Once armed with the real and synthetic networks examples, the rest of the course focuses on analysis of networks. The topics covered are not consistent from year to year. As this field is evolving so quickly, one of the goals for this course has been not to use a static book for teaching but rather to present slides and research articles based on exposure to ideas presented at the most recent conferences, such as NetSci (https://www.netsci2018.com/), CompleNet (http://complenet.weebly.com/), SIAM Workshop on Network Science (http://www.siam.org/meetings/ns18/), ASONAM (http://asonam.cpsc.ucalgary.ca/2018/). and Sunbelt (https://sunbelt.sites.uu. nl/). Before coming to my lectures, students have to watch TED talks on the topics that I will be teaching that day. This way, they have a different point of view of why the topic is interesting, and how other researchers have used it. They then come to my class with questions for me, which allows them to hear the answers I have on what I teach that day. The presentation slides are updated regularly, exposing the students to updated information and from several sources. It is important to me that students get multiple points of view on the topics, since network science researchers come from different fields of studies, emphasising the motivation and application they found. The slides are complimented with articles and recorded lectures from conferences and one or two standard books for references.

The rest of this chapter is organized as follows. Section 2 presents the course learning outcomes and objectives, followed by the content and software used in the class and detailed in Section 3. Section 4 gives the overview of the course format, detailed in Section 5, as well as including the assessment for each activity. Finally there is a conclusion and student feedback.

#### 2 Course Learning Outcomes and Objectives

The goal for students in this class is to develop the mathematical sophistication needed to apply learned methodologies to, and understand properties of new networks. This course enables them to have enough exposure and practice to readily use existing concepts or further read and understand published research as needed for future projects.

To achieve this, students analyze their personal network for the Network Profile Summary by practicing the introduced concepts of complex network analysis and by describing the structure of the chosen network. Furthermore, they contrast network models to real networks, by explaining features some complex networks have that others do not. This allows them to synthesize the new research in this evolving area and critique a peer's research. Students also read papers for which they have to grasp and explain new research ideas in complex networks.

The outcome of the course is that through new network research, the students will design new network models building on existing ones and available data. They will be able to design experiments to test hypotheses based on analyzed data and generate new methodologies by expanding on the designed experiments.

The learning outcomes above are achieved through building and analyzing personalized networks, reading scientific papers, writing technical research articles for publication, and presenting them at network science conferences. In my view, these are exactly the puzzle pieces for attaining the learning objective of the course: understanding the concepts, models, and methodologies needed to identify how to use knowledge of complex networks to produce a research article or apply in a realworld situation. This gives students the mathematical sophistication and confidence to use gained experience as situations arise.

#### **3** Course Content and Software

The course materials are available at http://faculty.nps.edu/rgera/MA4404.html. The topics of the course are the following:

- 1. Types of networks:
- Synthetic network models: Erdős–Rényi random networks, Watts-Strogatz small-world networks, the Barabási–Albert preferential attachment growing model and its variants, the Malloy-Reed configuration model, and the random geometric model;
- 3. Network statistics/properties: degree, clustering coefficient, diameter, density, shortest paths, node similarity, and homophily; and
- 4. Centralities: degree, closeness, betweenness, eigenvector, Katz, PageRank, hubs, and authorities.

These topics get augmented by presentations based on information from current conferences. Because of the fragmented literature–with inconsistent terminology and frequent reinvention of concepts and methodologies of network science due to the mix of the backgrounds of their researchers–this class builds on several manuscripts and conference presentations. Presentation slides are available for each lecture day at http://faculty.nps.edu/rgera/MA4404.html, and they are updated based on new research and information and animations from conferences. The main articles used as the class references are (a) Newman's 2003 article "The Structure and Function of complex networks" [48] which can be found at http://epubs.siam.org/doi/pdf/10.1137/S003614450342480 from SIAM and (b) Barabási–Albert's free and interactive book *Network Science* which can be found on his website at http://barabasi.com/networksciencebook (and also in print [8]).

The main visualization tool is the open-source graph visualization and manipulation software Gephi, found at http://gephi.org/ [1]. To complement Gephi's analysis ability, I generally use Python or R, open-source programming languages with wide interoperability, and other tools [10].

For Python users https://www.python.org/ [2], I suggest NetworkX (https://networkx.github.io/) and igraph (http://igraph.org/redirect.html), two Python libraries developed for the study of graphs and networks.

For R users https://www.r-project.org/ [3], I suggest igraph, http://kateto.net/ networks-r-igraph and an overview found here (http://www.necsi.edu/events/iccs6/ papers/c1602a3c126ba822d0bc4293371c.pdf) [24], or Statnet with the following tutorial from a Sunbelt conference: https://statnet.org/trac/raw-attachment/wiki/ Resources/introToSNAinRsunbelt2012tutorial.pdf.

Available tools are updated on a regular basis, as this information is not static: http://faculty.nps.edu/rgera/MA4404.html

#### 4 Course Format

Classes meet Monday through Thursday, for 50 min each. For the first 3 weeks of the quarter, interactive lectures are provided for each class. During this time, students are exposed to an overview of network science and real and synthetic networks.

There are two assignments due in these 3 weeks. The first assignment is the Introduction to Multilayer Networks Project, in which students are exposed to multilayer networks. The type of multilayer network that I am interested in captures each of the diverse types of relationships between the nodes into a separate layer of the network.

An example of such a network is the terrorist network in Fig. 1. I also provide the link to the comprehensive review (including temporal networks, networks of networks, and interdependent networks) which can be found in [16]. For visualization for multilayer networks, I suggest that they try Pymnet [44] found at http://people. maths.ox.ac.uk/kivela/mlnlibrary/, Muxviz [26] found at http://muxviz.net/, or Gephi [10] found at http://gephi.org/.

The students' assignment is to create a classroom multilayer network whose nodes are the students in the current class. Different student attributes get captured, which allow the formation of edges/relationships between the students, and are categorized into layers of a multilayer network. The students decide on the relationships they wish to capture, the end goal being to partition the class into teams of 3–4 students to write a research paper together. An expansion of this will follow in Section 5.2.

The second assignment for this period of time is to create or search for a network to analyze during the course of study, called the Network Profile Summary (Sect. 5.4). Each of these networks serves as data that each student analyzes for his/her homework by applying his/her understanding of the topics introduced in class that week. Thus the assignments are personalized, as the student chose the data, while all the students try all the learned concepts of the week and report their observations into one presentation slide per week (PowerPoint or  $L^{A}T_{E}X$ ). The requirement of summarizing their findings in one slide provides the opportunity for the student to



Fig. 1 An example of multiplex networks, in which each layer captures different relationships, such as friends, training, classmates, meetings, and operations

present synthesized information. Presenting the observed results enables the student to identify and explain the "why" behind the "what" of the findings, rather than asking the authority, the professor. Each student individually gives a 5-min presentation on his/her Network Profile Summary, based on the topics learned on that particular week. Presentations are followed by in-class discussion of that week's topics on a variety of networks that students present in class. An extended discussion on the Network Profile Summary follows in Sect. 5.4, and Gera et al. examine it in detail in [36].

Starting with the fourth week of the quarter is a transition to the following schedule:

- Mondays and Wednesdays: lectures.
- Tuesdays: teams meet to work on the Research Project (see Sect. 5.3). Each year a research topic is provided to the class, and each team finds a new methodology for solving one of the couple of choices of the open problems. While teams work on their projects, I work with each team to validate the direction of the research approach and to answer questions.
- Thursdays: students give their presentations on the Network Profile Summary (Sect. 5.4), followed by team discussions contrasting the results presented.

## 5 Student Learning and Assessment

The point value of the class activities are summarized in Table 1, with a longer discussion on each activity following the table.

# 5.1 In-Class Participation (70 Points)

The interactive teaching style requires everyone to participate in classroom discussions. Students are encouraged to be engaged in these discussions while giving everyone else a chance to confirm their understandings or mend their confusions. The in-class conversations allow students to modify and improve existing perceptions about the network science topics.

Table 1         The breakdown of	Activity	Points (of 300 total)	
points for the final grade	In-class participation	70	
	Introduction to Multilayer Networks Project	30	
	Network Profile Summary	100	
	Research Project	100	

Assessments of learning: Participation is measured by evidence of class preparation, interactions during class. This is objectively measured by asking relevant questions, showing the ability to express critical thinking, and making connections even if they are not correct. These behaviors show whether students are actively engaged or passively listening.

#### 5.2 Introduction to Multilayer Networks Project (30 Points)

The following has been used as an introductory project, modifying it for different cohorts of students, and it works well each time. While I talk about edge coloring in graph theory, there is a different purpose for the categories that form the colors there versus the layers; and thus this is the first exposure to multilayer networks.

As students will work in teams for the main Research Project detailed in Sect. 5.3, the first activity of this course provides the teams for the Research Project. That is, while learning about multilayer networks, students produce a multilayer network of their current class and identify a possible breakdown into teams to complete the Research Project. This way, while students start to think of multilayer networks for the first time, they have an interest in listening to the various solutions since (1) they thought about the problem as the whole class has the same task, and (2) they will be affected by the created teams.

Students work based on a description provided in advance and detailed below. Then they present the Introduction to Multilayer Network Project results at the end of the third week of classes. Each team has 10–13 minutes to convince me that their proposed teams (and reasoning for the team formation) should be the one to be adopted for the classroom. The following summarizes the project as given to the students:

Description: Research is a major component of this course. Since literature shows the best research is done in teams whose members have diverse backgrounds to integrate the research endeavor [40, 50], students are tasked to partition the class so that each team can accomplish the research goals. The students are provided with the class roster for MA4404 and some attributes that help them decide what relationships to add between the students. Such an example is shown in Table 2.

Goal: Students are asked to create a multilayer network (each attribute is captured in a different layer) and partition the MA4404 class into research teams. Ideally, members should complement each other based on the given attributes, and additional information can be collected. The goal is to minimize variability among performance of the teams, rather than to maximize the performance of one team.

Data: Table 2 presents sample data format used to create the class network, which could be augmented by other characteristics as the class sees useful. Each row represents a student, and each column captures the entry of that attribute per student.

Name	Major (dual degree with)	Known coding language (first/second), Beginner (B) / Intermediate (I)/ Advanced (A)	Military service	Graduating month/year	Other relevant skill(s)	Previous partners
Student 1	MA	Python (B)	Army	June '17	Good speaker	Student 1
Student 2	CS	Python (A) and R (I)	Navy	Sep '17	Good writer	Student 4, 8
	SE	None	Air force	Dec '17	Visual	
Student n	OR	R (A)	Marines	Sep '17	Detailed	

 Table 2 Possible data for the Introduction to Multilayer Networks Project

Tasks:

- Describe the methodology of network creation: Students must identify nodes, edges, and layers. Visualization is optional as they use their creativity to explain the network.
- Describe the methodology of team assignment: Students must describe the methodology for team creation, identify what characteristics were most important in selecting individuals for teams, and reasons why.
- Describe the results: Students must present their proposed teams and argument for why this distribution of talent meets the goals established for this project. Multiple solutions may be presented if needed, but not encouraged.
- Present conclusions and future work: Students expand on what they took away from this task. They provide other attributes they believe could inform these results and explain why those attributes matter.
- Assess learning: The assessment will be driven by the accuracy and creativity of the model and its solution. However, the following should shape the presentations and be taken into account: helpful visual aids and clear, complete, and organized presentation (labeled pictures/tables).

# 5.3 The Research Project (100 Points)

The goal is to explore a novel research topic and learn the process of turning an exploration into a research paper. To begin this process, each student will review one paper that will be used as one of the references for the project and turn in one to two paragraphs (or up to a page) synthesizing the paper. This allows students to critically look at the research, carefully analyzing examples of papers to follow as they work on their project.

I provide an environment in which students have a chance to think creatively and make educated hypotheses. As researchers learn a great deal from both success and failures, I maintain the perspective that mistakes are inevitable, and progress still happens since the missteps spark creativity and deepen understanding. While negative results are not always desirable, they will not impact the final grade if the procedure to obtain the hypothesis is correct. The final goals for the project are:

- A short, 5-minute update of the team research idea and findings each Tuesday.
- A final 20-minute presentation during the last week of classes.
- A team research paper (about 10 pages), due to the week of finals.

In 2017, the topic for the Research Project was to create a mathematical model for synthetic multilayer networks, as will be described next.

Creating multilayer synthetic networks (or generative models): We live in a connected world, where networks dominate our economy, our environment, and our society. Understanding these networks can aid researchers in devising plans for devastating natural disasters, such as the eruption of the Eyjafjallajökull volcano in 2010 [56] or the Ebola outbreak [42]. While real networks are insightful, they come with challenges: they are usually hard to obtain (such as repetitive samples of the same type of network) to create temporal networks; data collected to create networks may contain personally identifiable information, or the sampled data may be at the wrong scale; or it can be very time-consuming to create several data sets for analysis. Thus, researchers desire methodologies to create synthetic networks that mimic the real ones and that allow the researchers to change the parameters to create different scales of networks that have similar properties to those observed in the real networks.

Goal: The goal of this project is to create networks that are multilayer, have a varying parameter to get different scales, and have similar scaled properties to real ones at the layer and global levels (matching the properties of the real one when the synthetic network is at the same scale as the real network).

Data: For a multilayer data set, the following link for the European Union Airline Data can be used as an example: http://faculty.nps.edu/rgera/MA4404/EUAirports. zip. Students were encouraged to search other data sets as well. Larger data sets are available on Clauset's website [21] at https://icon.colorado.edu/.

A multilayer network has two or more layers based on the type of edges (See Fig. 1). A longer discussion can be found on Domenico's website [30] at https:// comunelab.fbk.eu/multinet.php or in Kivela et al.'s (2014) paper [43]. The choices for visualization tools are MuxViz [26], Pymnet [44], or Gephi [10] used to visualize each layer individually, or anything like it.

Assessments of learning: The weekly updates serve as formative assessments in preparation of the summative ones. I strongly suggest to students that they build their PowerPoint and research paper weekly, so I can provide feedback. I use the rubric in Table 3 for assessing the presentations. The outline below guides the Research Project that will emerge as a paper:

- 1. 25 points for ongoing weekly progress.
- 2. 75 points for research paper: Abstract, 5; Related Work, 10; Methodology, 10; Ingenuity (or reasoning for the existing method), 15; Analysis, 30; Conclusion, 5.

Criteria	Task	Detailed Step (0–10 points)
Content	An analysis is performed	Correct analysis synthesizing learned concepts
		A (9–10 points) Relevant and clearly explained findings, insightful contextualization of findings, and thoughtful synthesis and interpretation of metrics
		B (8 points) Minor errors
		C (7 points) Significant errors D (6 points) Major conceptual errors
		<i>F</i> (0–5 <i>points</i> ) Little to no work of merit
Presentation	Results are presented	Clarity and style of graphics (could they be presented in a more significant way)
		A (9–10 points) Slide Deck Clear and succinct slides, correct spelling and mathematical notation, figures and tables are labeled and have captions consistent in tense and active voice, references are provided, thoughtful synthesis and interpretation of metrics <i>Conveying the information</i> Clear verbal explanation, correct use of terminology while explaining, clear and loud speaking
		<i>B</i> (8 points) Minor errors
		C (7 points) Significant errors D (6 points) Incoherent presentation
		<i>F</i> (0–5 points)
		Little to no work of ment

 Table 3
 Assessment rubric for slides and all the presentations, see Gera et al. Reference [36]

This paper's quality is given by knowledge integration:

- 1. Accuracy and vision: The modeling assumptions need to be appropriate, and the model needs to be checked against true network(s). The publication needs to give insight beyond a restatement of existing work and the exposition of the raw analysis of data; however, it should be related to existing work so that it has a place the current field of research.
- 2. Critical reasoning and exposition of relevant course material: Contrast the newly introduced methodology/parameter to existing ones. Present arguments for this methodology, and test statistics demonstrating competence with the content of complex networks. Explain connections to the real world and the observations/ implications of the found results.

- 3. Clarity: Students must communicate the problem and questions addressing the introduced methodology and approach, their insights, solutions, and remaining open questions. Students are asked to make their explanations concise by eliminating unnecessary verbiage.
- 4. Rigor and precision: The resulting paper must be mathematically precise (using proofs if such results are presented) and logical in its reasoning throughout. Any methodology used should be justified, and limitations or assumptions should be clarified.

The two main parts of the project are:

- Theory development: Students propose a theoretical direction and present reasons for the new methodology.
- Data set analysis: students must compare the networks chosen for the Research Project to the synthetic networks they create, with the goal of showing similarities and discrepancies. Previous years' work on the Research Project has materialized in the following publications [5, 7, 11, 19, 22, 23, 25, 37].

#### 5.4 Network Profile Summary (100 Points)

By the end of the third week of classes, each student must choose a network that (s) he is interested in understanding. For the remaining of the quarter, (s)he will be analyzing the chosen network and present results about it. This type of a project transitions from knowledge exposure to practice (without mimicking the instructor's particular example).

During and outside the class, students have the opportunity to experiment on their network, exploring the variety of topics as they present and as their intellectual curiosity inspires them. At the end of each week, each student creates one presentation slide (or two slides if absolutely needed) with the performed analysis of the topics presented that week, which is applied to his/her individual network.

Each Thursday four to five students present their findings, followed by a class discussion of all the networks' analysis (similarities and dissimilarities). Each student presents three times during the quarter, each being worth 20 points; the final presentation wraps up the takeaways from the study of the network and incorporates the personal updates provided weekly, and it is worth 40 points. The final presentation slides are due the last day of classes.

Data: Students search for data based on their interest. They can also bring the data they analyze for their theses, as they usually take this course during the time they work on the thesis projects (allowing students to incorporate network theory into their own research). I also provide scholarly sources and data that can be downloaded from a list compiled over the years http://faculty.nps.edu/rgera/MA4404/ NetworkProfileSummaryResources.html. One can also collect personal data from Facebook or LinkedIn (anonymized and not published); also based on hashtags, one can easily collect data using Netlytic [38]. Examples: Now included are some examples from the 2017 cohort presentations, with the approval of the students. These slides have not been published; they are duplicates of the slides presented during the regular Thursday individual Network Profile Summary classroom presentations.

Major (Maj) Daniel Funk created the Global Maritime Transportation network (Figs. 2 and 3). There are 120 "ports and chokepoint," and the edges were built based on data of ports' exports and imports.

The data originated from CIA World Factbook https://www.cia.gov/library/publications/the-world-factbook [32]. The data is separated into a sea layer and a



Fig. 2 An introduction of the Global Maritime Transportation System, by Maj Daniel Funk. (a) Highlighting the two types of nodes, (b) Highlighting the two layers



Fig. 3 The two layers of the Global Maritime Transportation System, by Maj Daniel Funk. (a) Centrality analysis, (b) Modularity and communities

road layer based on real travel distances (in nine nautical miles) on sea and land routes between the locations.

The data was collected from Bing Maps, Google Maps, and SeaRates https://www. Searates.com [4]. The PowerPoint slides present the results of modularity and community detection. Community detection partitions the network into groups by maximize modularity; it is one approach to studying communities in networks [49].

Captain (CPT) Brian Weaver analyzed the Storm of Swords data from Game of Thrones (Figs. 4 and 5). The data was collected and used in the article [12] to see who is the most degree central node.

E	POS SCH	IGRADUATE OOL	(	Ce	ntral	ities	
Top 1 1 2 3 4 5 6 7 8 9 10	0 Degree cer Tyrion Jon Sansa Robb Jaime Tywin Cersei Arya Catelyn Joffrey	ntrality nodes: 0.339623 0.245283 0.245283 0.235849 0.226415 0.207547 0.188679 0.179245 0.169811 0.169811	Based on Degree centrality, these people are connected to the most people. Not many conclusions can be drawn other than they are important based on the number of interactions they have with characters.	Top 1 2 3 4 5 6 7 8 9 10	10 Betweenn noder Jon Robert Tyrion Daenerys Robb Sansa Stannis Jaime Arya Tywin	ess centrality 52 0.229965 0.209452 0.197913 0.157203 0.126964 0.12672 0.1027 0.099943 0.079607 0.065538	Betweenness is the only measure where Tyrion is not number one. Jon win here because he is connected to many peopl that no one else is (he ha ventured beyond the wall which not many character have) He therefore lies the shortest path between nodes Daenaerys also appears uniquely here.
<u>Top</u> 1 2 3 4 5 6 7 8 9 10	10 Eigenvect node Tyrion Sansa Jaime Cersei Robb Joffrey Tywin Arya Robert Catelyn	State         Constrainty           0.336637         0.278782           0.273663         0.246448           0.246448         0.24856           0.224856         0.222931           0.192485         0.192485	The top characters here are found in King's Landing, who are connected to many other well-connected people. Jon, who wins in betweenness, does not make the list. His connections are not well connected/important.	Top 1 1 2 3 4 5 6 7 8 9	O Closeness or Tyrion Sansa Robert Robb Arya Jon Jaime Stannis Tywin Eddard	entrality nodes: 0.512077 0.509615 0.5 0.488479 0.486239 0.479638 0.479638 0.479638 0.479638 0.469027 0.46087	Unique to this list is #3, Robert, a guy who has been dead for 2 books now. Many people can 'think' of this character an therefore share an edge, no matter where they are located. This causes his average distance to other nodes to be small. Also, the Starks (Sansa, Robb,

b



# **Community Detection**



Fig. 4 The Storm of Swords of Game of Thrones network, by CPT Brian Weaver. (a) Centrality analysis, (b) Louvain community detection



Fig. 5 The Storm of Swords as a small world and its homophily, by CPT Brian Weaver. (a) Comparing to Watts–Strogatz and Erdos–Renyi, (b) Homophily and assortativity analysis

WWW.NPS.EDU

10

Brian augmented their analysis based on the topics studied in the class. The examples particularly present the Louvain community detection results, in which the Louvain method is one way to partition the network in community based on maximizing modularity [14]. Porter et al. provide a comprehensive article on other methods of community detection [51].

Lieutenant Commander (LCDR) Kevin Garcia presented the Tesla Superchargers network (Figs. 6 and 7). The data and distances between the Superchargers were



Fig. 6 An introduction of Tesla's Superchargers network, by LCDR Kevin Garcia. (a) A view based on the geolocation of the data, (b) Louvain community detection

-

а									
	-lin	NAVAL POSTGRADUATE SCHOOL	т	ēsl	a Network –	Netw	ork	Comparisor	า
-		Deper Distribution for Sarah		o e e e à de e		Ndonn)	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		
		Barabasi Albert Number of node Number of edges Average degree:	Model <u>s</u> : 259 : 1764 13.622		Erdos Renyi Mod Number of nodes: 2 Number of edges: 1 Average degree: 13.	del 259 707 182	<u>W</u> 1 1	Vatts Strogatz Mod Number of nodes: 25 Number of edges: 183 Average degree: 14.0	<u>del</u> 9 13 00
		Observation: Clearly logarithmic I Distribution	Degree		Erdos Renyi and Watt Degree Distribution, wi	Observ s Strogatz gra th the Watts :	vation phs end Strogati	<u>1</u> : ded up with similar looking z Model having a much sm	s all
信し					distribution cente www.nps.edu	red around th	ie avera	age degree distribution.	4
b					distribution cente www.nps.EDU	red around tr	ie avera	age degree distribution.	4
b	-Le	NAVAL POSTGRADUATE SCHOOL			Centrali	ty - Ov	verv	age degree distribution.	4
b		NAVAL POSTGEADUATE SCHOOL	rality	Be	Centrali etweenness Cen	ty - Ov	/er\	view <u>Degree Centra</u>	4 ality
b	-15	NAVAL POSTGRADUATE SCHOOL Closeness Cent Top 10 Node	rality 25:	Be	Centrali etweenness Cen Top 10 Node	ty - Ov <u>s:</u>	/er\	view Degree Centra Top 10 Node	4 ality 25:
b	1. 2.	NAVAL POSTGRADUATE Closeness Cent Top 10 Node Topeka, KS Hay, KS	rality 25: 0.1691 0.1669	<u>B</u> e	Centrali etweenness Cen <u>Top 10 Node</u> Columbia, MO Topeka, KS	ty - Ov <u>trality</u> <u>5:</u> 0.3389 0.3373	/er\ 1. 2.	view <u>Degree Centra</u> <u>Top 10 Node</u> W.Springfield, M/ W. Hartford, CT	4 <b>ality</b> 25: A 0.1240 0.1240
b	1. 2. 3.	Topeka, KS Columbia, MO	rality 25: 0.1691 0.1669 0.1661	<u>B</u> e	Centrali etweenness Cen <u>Top 10 Node</u> Columbia, MO Topeka, KS Limon, CO	ty - Ov <b>trality</b> 0.3389 0.3373 0.3168	1. 2. 3.	view <u>Degree Centra</u> <u>Top 10 Node</u> W.Springfield, M. W. Hartford, CT Albany, NY	4 <b>ality</b> 25: A 0.1240 0.1240 0.1240
b	1. 2. 3. 4.	Topeka, KS Columbia, MO	rality 25: 0.1691 0.1669 0.1661 0.1637	<u>B</u> e 1. 2. 3. 4.	Centrali etweenness Cen <u>Top 10 Node</u> Columbia, MO Topeka, KS Limon, CO Hays, KS	ty - Ov <b>trality</b> 0.3389 0.3373 0.3168 0.3089	1. 2. 3. 4.	view <u>Degree Centra</u> <u>Top 10 Node</u> W.Springfield, M, W. Hartford, CT Albany, NY Milford, CT	4 ality 25: A 0.1240 0.1240 0.1240 0.1201
b	1. 2. 3. 4. 5.	Topeka, KS Columbia, MO Limon, CO Effingham, IL	rality 25: 0.1691 0.1669 0.1661 0.1637 0.1626	<u>B</u> e 1. 2. 3. 4. 5.	Centrali etweenness Cen <u>Top 10 Node</u> Columbia, MO Topeka, KS Limon, CO Hays, KS Glen Spring, CO	ty - Ov trality <u>s:</u> 0.3389 0.3373 0.3168 0.3089 0.2554	/er\ 1. 2. 3. 4. 5.	view <u>Degree Centra</u> <u>Top 10 Node</u> W.Springfield, M. W. Hartford, CT Albany, NY Milford, CT Milford, CT	4 ality 25: A 0.1240 0.1240 0.1201 0.1201
b	1. 2. 3. 4. 5. 6.	Topeka, KS Hay, KS Columbia, MO Limon, CO Effingham, IL Normal, IL	rality 25: 0.1691 0.1669 0.1661 0.1637 0.1626 0.1612	<u>B</u> e 1. 2. 3. 4. 5. 6.	Centrali etweenness Cen <u>Top 10 Node</u> Columbia, MO Topeka, KS Limon, CO Hays, KS Glen Spring, CO Effingham, IL	ty - Ov trality 5: 0.3389 0.3373 0.3168 0.3089 0.2554 0.1581	1. 2. 3. 4. 5. 6.	view <u>Degree Centra</u> <u>Top 10 Node</u> W.Springfield, M. W. Hartford, CT Albany, NY Milford, CT Milford, CT Newburgh, NY	4 ality 25: A 0.1240 0.1240 0.1240 0.1201 0.1201 0.1201
b	1. 2. 3. 4. 5. 6. 7.	Topeka, KS Hay, KS Columbia, MO Limon, CO Effingham, IL Normal, IL Springfield, IL	rality 25: 0.1691 0.1669 0.1661 0.1637 0.1626 0.1612 0.1608	<b>Be</b> 1. 2. 3. 4. 5. 6. 7.	Centrali etweenness Cen <u>Top 10 Node</u> Columbia, MO Topeka, KS Limon, CO Hays, KS Glen Spring, CO Effingham, IL Price, UT	ty - Ov trality 5: 0.3389 0.3373 0.3168 0.3089 0.2554 0.1581 0.1568	/er\ 1. 2. 3. 4. 5. 6. 7.	view <u>Degree Centra</u> <u>Top 10 Node</u> W.Springfield, M, W. Hartford, CT Albany, NY Milford, CT Milford, CT Newburgh, NY Darien, CT	4 ality 25: A 0.1240 0.1240 0.1240 0.1201 0.1201 0.1201 0.1201
b	1. 2. 3. 4. 5. 6. 7. 8.	Topeka, KS Hay, KS Columbia, MO Limon, CO Effingham, IL Normal, IL Springfield, IL Perry, OK	rality 25: 0.1691 0.1669 0.1661 0.1637 0.1626 0.1612 0.1608 0.1606	<b>Be</b> 1. 2. 3. 4. 5. 6. 7. 8.	Centrali etweenness Cen <u>Top 10 Node</u> Columbia, MO Topeka, KS Limon, CO Hays, KS Glen Spring, CO Effingham, IL Price, UT Albuquerque, NM	ty - Ov trality 5: 0.3389 0.3373 0.3168 0.3089 0.2554 0.1568 0.1568 0.1502	1. 2. 3. 4. 5. 6. 7. 8.	view <u>Degree Centra</u> <u>Top 10 Node</u> W.Springfield, M. W. Hartford, CT Albany, NY Milford, CT Milford, CT Newburgh, NY Darien, CT Darien, CT	4 ality 25: A 0.1240 0.1240 0.1201 0.1201 0.1201 0.1201 0.1201
b	1. 2. 3. 4. 5. 6. 7. 8. 9.	Topeka, KS Hay, KS Columbia, MO Limon, CO Effingham, IL Normal, IL Springfield, IL Perry, OK Salina, KS	rality 0.1691 0.1669 0.1661 0.1637 0.1626 0.1612 0.1608 0.1606 0.1606 0.1596	<b>Be</b> 1. 2. 3. 4. 5. 6. 7. 8. 9.	Columbia, MO Top 10 Node Columbia, MO Topeka, KS Limon, CO Hays, KS Glen Spring, CO Effingham, IL Price, UT Albuquerque, NM Holbrook, AZ	ty - O trality 5: 0.3389 0.3373 0.3168 0.3089 0.2554 0.1568 0.1568 0.1502 0.1345	1. 2. 3. 4. 5. 6. 7. 8. 9.	view <u>Degree Centra</u> <u>Top 10 Node</u> W.Springfield, M. W. Hartford, CT Albany, NY Milford, CT Newburgh, NY Darien, CT Darien, CT Greenwich, CT	4 ality 25: A 0.1240 0.1240 0.1201 0.1201 0.1201 0.1201 0.1201 0.1201 0.1201

WWW.NPS.EDU

Fig. 7 Centralities and comparable synthetic networks, by LCDR Kevin Garcia. (a) Highlighting the different types of nodes, (b) Highlighting the different types of nodes

compiled by Maj Daniel Funk, using Tesla Supercharger locations obtained from www.Tesla.com.

The network has 259 nodes and 1700 edges based on the distance travelled on a full battery assumed to be 250 mile range. The actual range for different Tesla models varies from 240 miles up to 335 miles. Edges connect charging stations of distances up to 250 miles apart.

Kevin noted that, "While many networks slowly evolve over time, Tesla (and specifically Elon Musk) built the Tesla supercharger network because he was told that he wouldn't be able to. Instead of it evolving over time, he forced it, to make cross-country routes in a Tesla possible. This caused there to be very few nodes that are out on the periphery and not connected well to the graph. There are still several nodes that have a high degree in the metropolitan areas."

He concluded in his presentation slides [35] that for future work one should.... "analyze the network using all distances from each charging station to all other charging stations, not just the ones within 250 miles. From this information, edge weights could be added based on distances."

A more detailed explanation of the Network Profile Summary can be found in [36]. Some of previous students' work on the Network Profile Summary has become part of their theses publications [6, 34, 46, 54].

Assessment rubric for weekly presentations (and final presentation): This knowledge, skill, and ability assessment are both formative and summative assessment, allowing students to incorporate weekly feedback to refine the final project. A refined comprehensive presentation slide deck including the weekly draft presentation slide and a conclusion is due during the finals week. The details of the assessment are captured in Table 3.

#### 6 Conclusion and Feedback from Students

This course and the teaching methods practiced have one underlying principle in the design: allow students the freedom in choosing what they analyze, while the instructor provides minimum guidance and explanations of the observed phenomena on the networks that students analyze. The instructor presents the high level overview of the topics covering the "why study" (during lectures and using TED talks) and "how to do it". However, the students apply the concepts to their chosen network, presenting the reasoning for what is observed.

Much like a building must pass certain control thresholds and a certain minimum quality control before moving on to the next stage, there is a need for a process approach to teaching that allows all students to obtain an accepted/approved level of understanding of newly learned topics before exposing them to additional information. With the current method, I believe that student's understanding is supported to be above a certain threshold by increasing their interest (letting them choose the network they study) and by asking them to find the "why" behind the "what is observed." I strongly believe that while I can provide them answers (i.e. teaching them), they only hear and understand it if they ask the questions to which I provide the answers. The practices described in this article allow the students to ask the questions, in order for them to hear my answers (the messages I express through my teaching) or find their own answers (which certainly empowers them and makes the inforamtion relevant and interesting).

The structure of the course facilitates an environment in which students can learn at their pace and depth. Each student is in control of the depth (s)he goes to understand her/his network Network Profile Summary, and each student must at a minimum apply all the topics (s)he was exposed to each week. The students then present their analyses and conclusions of what they learned of their network based on the new concepts learned that week. This step-by-step exploration of the unknown allows the students and professor to have an incremental approach to navigating through their network. Furthermore, this learning style is unlike the traditional one in which students mimic the examples worked in the book or classroom. This promotes creativity and allows the student to decide how to synthesize the information rather than mimicking an existing template.

Secondly, students work on Research Projects as a team, as they most probably work on everything else in the military environment. This way everyone contributes based on their strength(s), they see how useful they are to the team, and they build up confidence. The fact that the topic given is open for research rather than being a solved problem, allows them to consider ideas beyond what people in the field would normally think of, as well as freeing them from trying to come up with the answer that the professor already had. It allows them to experience what it means to do research. Neither the graduation nor the grade depend on the result itself but rather on creativity and the use of critical thinking based on the newly learned concepts.

The third reason is that pursuing their individual, or team's, interests while learning network science enables the student to correlate the learned concepts and methodologies to their preexisting knowledge. This is because they have freedom in their actions while taking responsibility for the choices of data, methodology, and presented results. These choices make the class relevant and allow students to synthesize the information they are exposed to, which makes the topics accessible in the future as need arises.

There are multiple advantages and ramifications to allowing the students to practice concepts on their personal network used for the Network Profile Summary since there are no expectations of certain results, and nobody to compare against as is traditionally done with homework. The expectation is that they make sense for the network, or the student finds an explanation if they don't make sense. Students first and foremost feel empowered by taking responsibility for their own learning since they have to understand and explain the results. Also, they answer questions over the interpretation of the results, as they are the experts of their own network. This motivates them to search for the reasons behind what is observed and gives them confidence in their findings.

Using the Network Profile Summary gives them the potential to obtain the needed understanding of the concepts before moving on. This is based on the observation that junior teachers get a deeper understanding of topics when they have to teach them to others, trying to find the best way to explain the topic and explaining the reasons behind the presented information. Lastly, it allows the students to further present a synthesis of the results obtained into convincing, coherent, and cohesive arguments for their personal network and team project.

Following is the feedback from students in the course:

- "This class is structured in the way I thought graduate school classes should be structured. Open research questions were effective in inspiring more advanced learning. An excellent class."
- "Great course! I enjoyed exploring the course concepts by implementing them on my network profile. Great way to learn!"
- "This was an engaging and interactive course. It covered an incredibly interesting topic and the instructor did a great job bridging the 'math' with the real world applications. I enjoyed working in the team for projects and believe that this is by far the best way to learn. The lectures were interesting, improved my understanding of the material, and contributed directly to the quality of the projects. The instructor is passionate about the topics and passed that enthusiasm to the class. I wouldn't change anything."
- "Best. Teacher. Ever. Loved the course, and if you had a mic Ralucca, this is the part where you would drop it and walk away."

**Acknowledgments** I would like to thank the reviewers for their helpful and constructive comments that greatly improved the quality of the current manuscript. This material represents the position of the author and does not reflect the official policy or position of the US Government.

#### References

- 1. Gephi. http://gephi.org/. [Online; accessed 2017-12-31].
- 2. Python. https://www.python.org/. [Online; accessed 2017-12-31].
- 3. R software. https://www.r-project.org/. [Online; accessed 2017-12-31].
- 4. Searates.com. https://www.SeaRates.com. [Online; accessed 2017-12-31].
- Allain, R., Gera, R., Hall, R. and Mark Raffetto, (2016) M. Modeling network community evolution in YouTube comment posting. In International Conference on Social Computing, Behavioral-Cultural Modeling, & Prediction and Behavior Representation in Modeling and Simulation (BRIMS). http://sbp-brims.org/2016/proceedings/LB115.pdf.
- Allen, G. (2017) Constructing and classifying email networks from raw forensic images. Master's Thesis, 2016.
- Alonso, L., Crawford, B., Gera, R., House, J., Knuth, T., Méndez-Bermúdez, J. and Miller, R. (2017) Identifying network structure similarity using spectral graph theory. Applied Network Science,
- 8. Barabási, A-L. (2016) Network Science. Cambridge University Press,
- 9. Barabási Albert, R.(1999) Emergence of scaling in random networks. science, 286(5439):509–512,.
- Bastian, M, Heymann, S., Jacomy, M, et al. (2009) Gephi: an open source software for exploring and manipulating networks. Icwsm, 8:361–362,.
- Max Berest, Ralucca Gera, Zachary Lukens, N Martinez, and Ben McCaleb. Predicting network evolution through temporal twitter snapshots for Paris attacks of 2015. In International Conference on Social Computing, Behavioral-Cultural Modeling, & Prediction and Behavior Representation in Modeling and Simulation (BRIMS). http://sbp-brims.org/2016/proceedings/ LB111.pdf, 2016.
- 12. Andrew Beveridge and Jie Shan. Network of thrones. Math Horizons, 23(4):18-22, 2016.

- 13. John Biggs. What the student does: Teaching for enhanced learning. Higher education research & development, 18(1):57–75, 1999.
- 14. Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfold- ing of communities in large networks. Journal of statistical mechanics: theory and experiment, 2008(10):P10008, 2008.
- Phyllis C Blumenfeld, Elliot Soloway, Ronald W Marx, Joseph S Krajcik, Mark Guzdial, and Annemarie Palincsar. Motivating project-based learning: Sustaining the doing, supporting the learning. Educational psychologist, 26(3–4):369–398, 1991.
- 16. Stefano Boccaletti, Ginestra Bianconi, Regino Criado, Charo I Del Genio, Jesus Gomez-Gardenes, Miguel Romance, Irene Sendina-Nadal, Zhen Wang, and Massimiliano Zanin. The structure and dynamics of multilayer networks. Physics Reports, 544(1):1–122, 2014.
- 17. Charles C Bonwell and James A Eison. Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports. ERIC, 1991.
- Ann L Brown and Joseph C Campione. Guided discovery in a community of learners. The MIT Press, 1994.
- Samuel Chen, Joyati Debnath, Ralucca Gera, Brian Greunke, Nicholas Sharpe, and Scott Warnke. Discovering community structure using network sampling. In 32nd ISCA International Conference on Computers and Their Applications. Springer, 2017.
- 20. Ken Cherven. Mastering Gephi network visualization. Packt Publishing Ltd, 2015.
- 21. Aaron Clauset. Index of complex networks.
- 22. Brian Crawford, Ralucca Gera, Jeffrey House, Thomas Knuth, and Ryan Miller. Graph structure similarity using spectral graph theory. In International Workshop on Complex Networks and their Applications, pages 209–221. Springer, 2016.
- Brian Crawford, Ralucca Gera, Ryan Miller, and Bijesh Shrestha. Community evolution in multiplex layer aggregation. In Advances in Social Networks Analysis and Mining (ASONAM), 2016 IEEE/ACM, pages 1229–1237. IEEE, 2016.
- 24. Gabor Csardi and Tamas Nepusz. The igraph software package for complex network research. Inter- Journal, Complex Systems, 1695(5):1–9, 2006.
- Benjamin Davis, Ralucca Gera, Gary Lazzaro, Bing Yong Lim, and Erik C Rye. The marginal benefit of monitor placement on networks. In Complex Networks VII, pages 93–104. Springer, 2016.
- Manlio De Domenico, Mason A Porter, and Alex Arenas. Muxviz: a tool for multilayer analysis and visualization of networks. Journal of Complex Networks, 3(2):159–176, 2015.
- Ton De Jong and Wouter R Van Joolingen. Scientific discovery learning with computer simulations of conceptual domains. Review of educational research, 68(2):179–201, 1998.
- Edward L Deci, Robert J Vallerand, Luc G Pelletier, and Richard M Ryan. Motivation and education: The self-determination perspective. Educational psychologist, 26(3–4):325–346, 1991.
- Josep Diaz, Mathew D Penrose, Jordi Petit, and Maria Serna. Linear orderings of random geometric graphs. In Graph Theoretic Concepts in Computer Science, volume 1665, pages 291–302. Springer, 1999.
- 30. Manilo De Domenico. Multilayer networks.
- Leonhard Euler. Leonhard euler and the konigsberg bridges. Scientific American, 189(1):66– 70, 1953.
- CIA Factbook. The world factbook. See also: https://www.cia.gov/library/publications/theworld-factbook, 2010.
- 33. Scott Freeman, Sarah L Eddy, Miles McDonough, Michelle K Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary Pat Wenderoth. Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences, 111(23):8410–8415, 2014.
- Daniel Funk. Analysis of the global maritime transportation system and its resilience. Master's Thesis, 2017.
- 35. Kevin Garcia. The tesla's superchargers network. unpublished.

- 36. Ralucca Gera, Jessica M Libertini, Jonathan W Roginski, and Anthony Zupancic. The network profile summary: exploring network science through the lens of student motivation. Journal of Complex Networks, 2017.
- 37. Ralucca Gera, Ryan Miller, Miguel Miranda Lopez, Akrati Saxena, and Scott Warnke. Three is the answer: combining relationships to analyze multilayered terrorist networks. Advances in Social Networks Analysis and Mining (ASONAM), 2017.
- 38. A Gruzd. Netlytic: Software for automated text and social network analysis, 2016.
- 39. Aric Hagberg, Pieter Swart, and Daniel S Chult. Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Laboratory (LANL), 2008.
- Martin Hoegl and Hans Georg Gemuenden. Teamwork quality and the success of innovative projects: A theoretical concept and empirical evidence. Organization science, 12(4):435–449, 2001.
- 41. Paul A Kirschner, John Sweller, and Richard E Clark. Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educational psychologist, 41(2):75–86, 2006.
- 42. Maria A Kiskowski. A three-scale network model for the early growth dynamics of 2014 west Africa ebola epidemic. PLoS currents, 6, 2014.
- 43. Mikko Kivela, Alex Arenas, Marc Barthelemy, James P Gleeson, Yamir Moreno, and Mason A Porter. Multilayer networks. Journal of complex networks, 2(3):203–271, 2014.
- Mikko Kivela. Multilayer networks library for python (pymnet). http://people.maths.ox.ac.uk/ kivela/mln\_library/. [Online; accessed 2017-12-31].
- 45. Eric D Kolaczyk and Gabor Csardi. Statistical analysis of network data with R, volume 65. Springer, 2014.
- 46. Ryan Miller. Purpose-driven communities in multiplex networks: Thresholding user-engaged layer aggregation. Master's Thesis, 2016.
- 47. Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. Random structures & algorithms, 6(2–3):161–180, 1995.
- Mark EJ Newman. The structure and function of complex networks. SIAM review, 45(2):167– 256, 2003.
- Mark EJ Newman. Modularity and community structure in networks. Proceedings of the national academy of sciences, 103(23):8577–8582, 2006.
- 50. Max Nielsen-Pincus, Wayde Cameron Morse, Jo Ellen Force, and JD Wulfhorst. Bridges and barriers to developing and conducting interdisciplinary graduate-student team research. Ecology & Society, 2007.
- Mason A Porter, Jukka-Pekka Onnela, and Peter J Mucha. Communities in networks. Notices of the AMS, 56(9):1082–1097, 2009.
- 52. Dale H Schunk. Self-efficacy and academic motivation. Educational psychologist, 26(3–4):207–231, 1991.
- Jeffrey Travers and Stanley Milgram. The small world problem. Psychology Today, 1:61–67, 1967.
- Scott Warnke. Partial information community detection in a multilayered network. Master's Thesis, 2016.
- Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world' networks. nature, 393(6684):440–442, 1998.
- 56. Dunn S. Wilkinson, S.M. and S. Ma. The vulnerability of the European air traffic network to spatial hazards. Natural Hazards, 60(3):1027–1036, 2012.