2021 ASEE ANNUAL CONFERENCE
Virtual Meeting | July 26–29, 2021 | Pacific Daylight Time
Paper ID #34013

Gratitude and Graph Theory in the Time of Coronavirus

Prof. Gerald J. Wang, Carnegie Mellon University

Jerry Wang is an Assistant Professor of Civil and Environmental Engineering, and Mechanical Engineering (by courtesy) and Chemical Engineering (by courtesy), at Carnegie Mellon University. He received his BS in 2013 from Yale University (Mechanical Engineering, Mathematics and Physics), SM in 2015 from MIT (Mechanical Engineering), and PhD in 2019 from MIT (Mechanical Engineering and Computation). He performed postdoctoral research at MIT in Chemical Engineering. He is a member of the inaugural cohort of the Provost's Inclusive Teaching Fellowship at CMU, was the 2020 recipient of the Frederick A. Howes Scholar Award in Computational Science and the 2016 MIT Graduate Teaching Award in the School of Engineering, and is an alumnus of the Department of Energy Computational Science Graduate Fellowship and the Tau Beta Pi Graduate Fellowship.

Wang directs the Mechanics of Materials via Molecular and Multiscale Methods Laboratory (M5 Lab) at CMU, which focuses on computational micro- and nanoscale mechanics of fluids, soft matter, and active matter, with applications in Civil and Environmental Engineering across the nexus of water, energy, sustainable materials, and urban livability. The M5 Lab is particularly interested in particle-based simulations, systems out of equilibrium, uncertainty quantification in molecular simulations, and high-performance computing. He teaches courses in molecular simulation and computational/data science.

Gratitude and Graph Theory in the Time of Coronavirus

Abstract

In this work, we describe two activities that we have designed – suitable for an introductory undergraduate course in computational and data science – that introduce rudimentary principles of graph theory. Both activities feature simple premises (yet have the potential for significant depth, depending on student interest and mathematical maturity), strengthen students' abilities in mathematical and computational reasoning, and interweave timely "out-of-the-classroom" themes that will resonate in 2021 and beyond. We also provide qualitative results from the deployment of these activities in a sophomore-level civil and environmental engineering class.

Introduction

Graph theory is the study of collections of objects, in which any pair of objects may or may not be interconnected. A longtime favorite of mathematicians and computer scientists, graph theory is an important conceptual framework for all civil and environmental engineers (CEEs) to be fluent in, as it is the natural language for describing networks, ranging from traffic systems to pollution transport pathways to a host of applications in computational mechanics and beyond. This importance will only grow with time, as we enter an increasingly data-rich and computer-aided CEE landscape, which requires graph-based approaches for organizing and analyzing complex and interconnected data streams. Aside from its relevance to numerous CEE applications, graph theory also affords prime opportunities to reinforce and enrich students' proficiency with linear algebra, part of the bedrock of an engineer's mathematical training.

In this work, we describe two educational activities that serve as gentle forays into graph theory for an undergraduate CEE audience. Based on our positive experience with these activities – and given the ever-increasing importance of this subject – we believe that there is a compelling case for introducing graph theory as early as possible in a computing-forward undergraduate CEE curriculum.

Educational Context

The activities described in this work were developed for a course entitled, "Intro to Computational and Data Science in Civil and Environmental Engineering." It is a required course in the Carnegie Mellon University (CMU) CEE undergraduate program, generally taken by students in their sophomore year, with a typical enrollment between 30 and 45 students. This course presumes prior knowledge of computer programming fundamentals and (single-variable) calculus, but does not assume prior exposure to more advanced mathematics (including linear algebra, which is of particular relevance to the activities described here). Although most students (> 90%) have some prior experience with vectors (including from first-year physics) and with data structured as arrays (including from a first-year computing course), many students have not formally worked with matrices before this course. This course has been taught by CEE faculty ever since it was first created in the late 1990s (when it was first created, it was entitled, "Intro to Computer Applications in Civil and Environmental Engineering").

Computing and data science play critical roles in the CEE undergraduate (and graduate) curriculum at CMU. The undergraduate curriculum in our department provides students with a grounding in traditional CEE material, but has a particular emphasis on empowering students to play an active role in reimagining the field of CEE in the future. This course establishes the foundation for further computing (and sensing) skill development in required junior- and senior-level lab and project courses, including our senior capstone design course. Before graduation, a significant number of undergraduates also elect to take at least one graduate-level course with a strong computational focus.

Both educational activities described in this work were developed and deployed in the Spring 2020 semester, after the transition from in-person instruction to remote instruction. In particular, the first activity ("Graph Theory and Disease Transmission") was released five days after the State of Pennsylvania issued its first "Stay at Home" order; the second activity ("Graph Theory and Gratitude") was carried out over a two-week period at the end of the Spring 2020 semester.

An (Infinitesimal) Glossary of Relevant Graph Theory Vocabulary

We begin by very briefly introducing graph theory vocabulary relevant to the discussion below; as we could not possibly begin to do justice to this enormous subject in a single paragraph, we refer the interested reader to one of many introductory texts on the matter [1].

A **graph** is a mathematical structure that features a collection of **nodes** (each representing some object), some pairs of which are connected via **edges**. A graph may be **undirected** (in which case each edge simply establishes a symmetric connection between two nodes) or **directed** (in which case an edge also designates some asymmetry between its terminal nodes, e.g., a flow *from* one node *to* the other). There is a natural correspondence between graphs and matrices. In particular, an **adjacency matrix** is filled with elements that indicate whether any given pair of nodes are connected by an edge. A graph may be **connected** (in which case there exists a path from every node to every other node, formed by traversing edges); otherwise, we call the graph **disconnected**.[1] We term a disconnected graph with two individually connected components as a **bipartite** graph.

[1] In the case of directed graphs, depending on whether paths follow the specified flow directions along each edge, one may further distinguish between **weak connectivity** (in which case all nodes are connected only if directed edges are replaced with undirected edges), **semi-connectivity** (in which case a directed-edge-respecting path exists between all pairs of nodes, but not necessarily in both directions), and **strong connectivity** (in which case a directed-edge-respecting path exists between all pairs of nodes, in both directions). We draw these delineations for the sake of completeness given the ideas presented; for the purposes of the activities described below, these distinctions are immaterial.

"Graph Theory and Disease Transmission"

This activity served as the majority component of a one-week collaborative homework assignment. In this activity, students are asked to develop and perform Monte Carlo simulations in order to model disease transmission on a graph, in which nodes represent individuals and edges represent social connections, which may change over time depending upon individuals' social practices (additional details in Appendix A). Students assemble and interpret an adjacency matrix that represents a social network (Figure 1), and are encouraged to develop additional features including, e.g., contagiousness and social distancing. Students are tasked with working in teams to independently generate open-ended questions that can be studied using these simulations, and with writing a report to synthesize their findings.

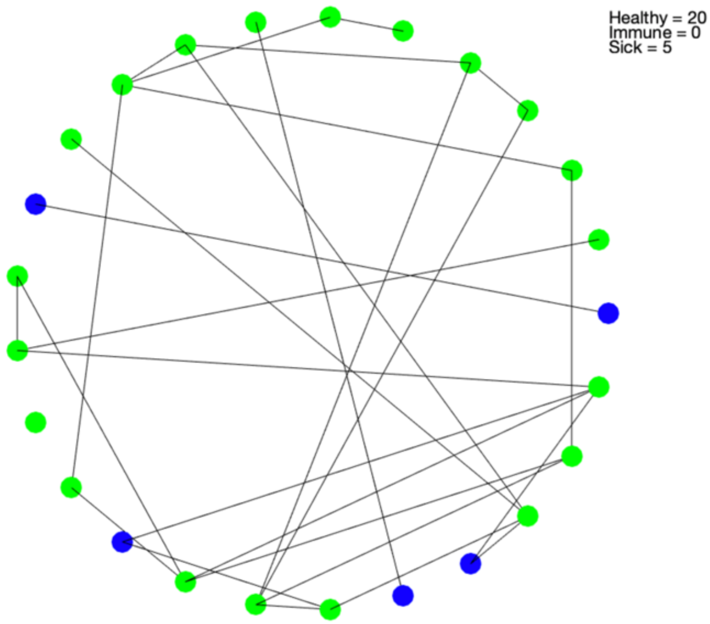

Figure 1: Visualization of a graph produced as part of "Graph Theory and Disease Transmission" in Spring 2020. Green nodes represent healthy individuals, blue nodes represent sick individuals, and edges represent social connections between pairs of individuals.

This activity reinforces numerous learning objectives. First and foremost, this activity provides ample opportunities to practice the fundamentals of computer programming, linear algebra, data visualization, and technical communication, all within a collaborative context. Of particular technical interest, this work introduces the physical interpretation of a matrix-vector product as an evolution operation for a (linear) dynamical system. For many students, this (short) independent research activity also serves as an accessible, relatively low-stakes introduction to the challenges and excitement of original scientific research. Finally, this activity encourages students to practice making explicit connections between mathematical concepts, real-world engineering problems, and policy.

"Graph Theory and Gratitude"

In the second activity, designed for the closing weeks of a course, students are invited to write notes of appreciation, anonymous or otherwise, for peers who have elevated their learning experience during the semester (additional details in Appendix B). These notes are collected and distributed to their intended recipients; based on the flow of notes, the instructor constructs an anonymized directed graph (Figure 2).

This activity reinforces numerous learning objectives. At a basic level, by presenting yet another context in which a data set is naturally encoded within a graph, this activity reinforces technical material covered several weeks earlier in the course. Moreover, this activity is an avenue to expose especially mathematically curious students to more-advanced topics near the end of term.[2] Although these topics are generally beyond the scope of this course, this graph was used to introduce the principles of cyclicity, partitioning, and spectral analysis in office hours discussions with interested students. This activity also produces a striking visual representation of the interconnectedness of a vibrant classroom, highlighting the sheer number of interpersonal relationships forged through teamwork and social connection. Such reminders are important in an in-person educational context and are even more important for students participating in remote education, which (in a vacuum) has a perverse tendency to nucleate and exacerbate feelings of isolation. As has been consistently demonstrated in the education literature, a healthy social and emotional climate is critical for students to succeed in the classroom [2]. This activity also centers within students an ethic of conscientious gratitude, which is well aligned with curricular goals that emphasize teamwork, socially conscious engineering, and lifelong learning.

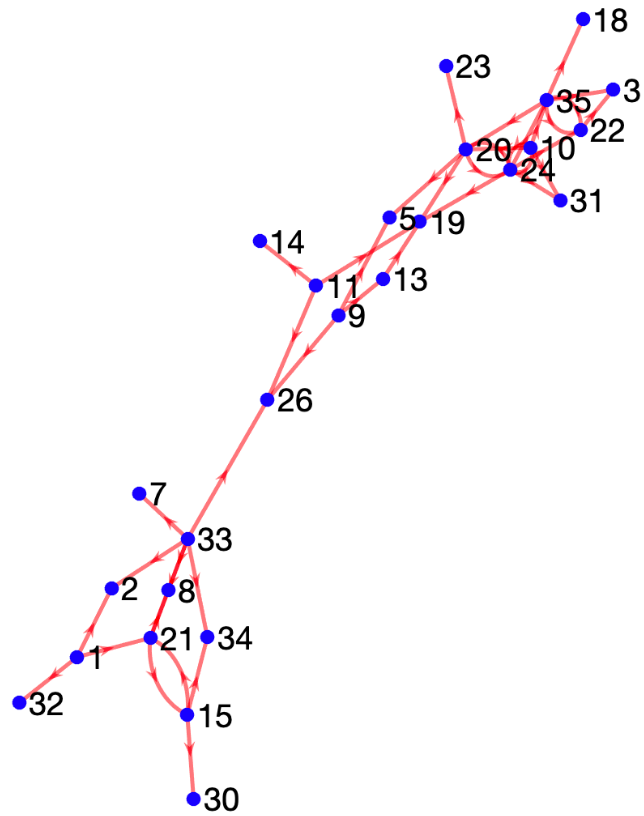

*Figure 2: Directed graph showing "flows" of gratitude between (anonymized) members of "Intro to Computational and Data Science in CEE" in Spring 2020. Nodes represent students and (directed) edges represent a note of gratitude submitted as part of this activity; the direction of each arrow indicates gratitude expressed **from** one student **to** another student. To maintain anonymity, each student in the course is assigned a random integer before the graph is constructed.*

This activity also generates an auxiliary source of data that reveals patterns of intramural collaboration and supportiveness that may not be immediately obvious to the instructional staff, especially in remote education. This data can supplement primary assessments of course participation; critically, this data facilitates more-equitable assessment of class engagement, especially for students who are less inclined to participate in synchronous classroom discussions but are nevertheless enthusiastic to study with peers.

---

[2] We believe that such avenues are particularly important in required engineering courses, which generally assume a reasonably high level of mathematical aptitude, but often feature a broad distribution in enthusiasm for more advanced mathematical material.

Discussion

In Spring 2020, a total of 12 reports were submitted for "Graph Theory and Disease Transmission"; these reports included contributions from all 38 students in the course (100% assignment submission rate). The high rate of engagement with "Graph Theory and Gratitude" was also notable: Despite being a strictly optional activity, nearly 50% of all course participants submitted notes of gratitude, and approximately 90% of all students in the course were mentioned within these notes.

The reports for the former activity contained numerous prescient insights for managing COVID-19 transmission risk. Without specific prompting, five of the 12 reports contained elaborate discussion (and in some cases detailed qualitative and quantitative guidelines) for forming "social pods," a practice that is now widespread not only on college campuses but throughout society [3]. Two of the reports focused on quantitatively modeling the impact of "short but massive social mixing" events; both reports supplied evidence that such events can dramatically accelerate a pandemic, and one report in particular argued that typical levels of social exposure associated with graduation ceremonies would lead to unacceptable levels of disease spread. One team's report featured an extensive effort to accelerate the performance of the simulation code, inspiring the members of this team to independently study computing topics beyond the scope of this sophomore-level course.

The notes produced through the latter activity also generated several non-trivial (and, in a few cases, remarkable) insights. Although most notes corresponded to collaborations that were already clear to the instructional staff, several notes revealed previously unknown collaborations, including a few that were formed only after the switch to remote instruction. In one memorable note, a student explained that the routine and energetic participation of a classmate motivated this student to become a more active class participant. Several students remarked upon the significant mental health benefits they enjoyed by working with various peers. These notes showed that certain individuals were particularly helpful to their classmates in particular sections of the course (e.g., two students were praised by several classmates for their insightfulness in the specific area of numerical optimization). Of particular curiosity was that this activity revealed a (nearly) bipartite structure within the class (in fact, the graph in Figure 2 would be strictly bipartite but for Node 26). This information was the basis of a recommendation to instructors of a junior-level project course to assign project teams that promoted mixing between these (nearly) disjoint sets.

Both activities received significant amounts of positive feedback. Much of this feedback was provided through student discussions with instructional staff in office hours; approximately 10% of course participants also referred to these activities in their written end-of-term course evaluations, with particular positive emphasis on their degree of customization to the unusual trajectory of the semester. It is also worth mentioning that following the conclusion of the course, specific and remarkable anecdotes about intellectual generosity and camaraderie from "Graph Theory and Gratitude" served as strong supporting material for several letters of reference.

Although an uneasy subject, we also find it worth discussing the relationship between current-events-motivated activities and a phenomenon that has long pervaded college campuses (and that has ballooned to new heights during the pandemic): Cheating and related violations of academic integrity [4]. It is well known to educators that there are a plethora of online sources for illicitly obtaining homework solutions.[3] Although the goal of this work is not to offer prescriptive solutions for reducing this (enormous) problem, we offer the following (anecdotal but meaningful) observations: (1) The week during which the "Graph Theory and Disease Transmission" activity was carried out featured the single highest amount of student participation in office hours, by a wide margin (nearly 60% of all students attended virtual office hours or sought out supplemental assistance at least once that week) and (2) In grading the submissions for this assignment, the instructional staff did not identify a single academic integrity concern throughout the entire class. Beyond the obvious interpretation (namely, that it is impossible to seek pre-existing solutions for a novel assignment), we believe that these observations support another meaningful conclusion: The development of timely and teamwork-driven activities can increase students' genuine interest about, and engagement with, challenging technical material. This observation is a classic manifestation of the well-understood phenomenon that increasing student motivation, especially by providing authentic tasks grounded in topics of student interest, leads to higher quality of learning [2].

Conclusions and Future Directions

We have described two activities – "Graph Theory and Disease Transmission" and "Graph Theory and Gratitude" – that are suitable for a sophomore-level CEE course on computing and data science, and that build and reinforce knowledge about graph theory. Both activities are closely tied to core engineering learning objectives, and have the potential to yield numerous ancillary benefits, including unusually high levels of course engagement, novel and valuable insights into social networks, and data to support decision-making in end-of-semester grading.

Based upon the success of "Graph Theory and Disease Transmission," the instructional staff for this course is currently developing additional short (single week) research activities to use in future offerings of the course. In the Fall 2020 semester, "Graph Theory and Gratitude" was successfully adapted for use in a graduate course on molecular simulation, including discussions of the resulting graph at a much higher level of mathematical complexity, and using concepts developed within the course.

Ultimately, graphs are only intellectually vibrant structures because of their embedded connections; the exact same is true of classroom communities. In the time of coronavirus, and with a final nod to Gabriel García Márquez[4], we are optimistic that these exercises can help a class discover (and with great delight) that one does not appreciate one's classmates just because they are one's classmates, but rather because of the friendships formed while learning together.

---

[3] The claim by the education technology company Chegg that their products (including solution manuals for textbook problems and crowd-sourced solutions for assignments at schools around the world) are expressly for helping students to "learn and understand" [4] rings as hollow as Captain Renault's "shock" in *Casablanca* [5].

[4] The 1982 Nobel Laureate in Literature, and perhaps best known for his novel *Love in the Time of Cholera*, a personal favorite of this author and the inspiration for the title of this manuscript.

Acknowledgements

G.J.W. gratefully and enthusiastically acknowledges the support and community of the CMU Provost's Inclusive Teaching Fellowship.

References

[1] R. J. Wilson, *Introduction to Graph Theory*. USA: John Wiley & Sons, Inc., 1986.
[2] S. A. Ambrose, M. W. Bridges, M. DiPietro, M. C. Lovett, and M. K. Norman, *How learning works: Seven research-based principles for smart teaching.* San Francisco, CA, US: Jossey-Bass, 2010.
[3] M. W. Moyer, "The Dos and Don'ts of 'Quarantine Pods,'" *New York Times*, 2020.
[4] J. R. Young, "With No Study Buddies, More College Students Turn to Cheating," *EdSurge*, 2020. [Online]. Available: https://www.edsurge.com/news/2020-10-06-with-no-study-buddies-more-college-students-turn-to-cheating.
[5] J. Rosenthal, "Shocked!," *New York Times*, 1997.

Appendix A: Prompt for "Graph Theory and Disease Transmission"

On our course webpage, I have provided you with a very simplistic (but nevertheless insightful) simulation tool. This code is by no means efficient (I have developed this tool in about 1.5 hours, and I have deliberately used mostly (inefficient) for loops instead of (efficient) vectorization in order to reduce the time it takes for you to interpret the crucial functions... you are welcome to vectorize any parts that you feel inspired to tackle). However, this code is more than good enough to develop an understanding of the roles played by population size, initial fraction of sick individuals, social distancing, ease of recovery, and immunity, in the trajectory of a pandemic.

The backbone of this simulation tool is linear algebra. For example, the disease transmission step is simply a matrix-vector product; the implementation of immunity is a vector-vector element-wise product. This tool is also, at its core, a Monte Carlo simulation since it makes use of a randomly generated set of sick individuals and a randomly generated contact network.

Our primary goal is to use this code to study disease transmission (and potentially make qualitative policy suggestions based on the insights we develop). Our secondary goal is to get you warmed up for the final projects: I am giving you all a LOT of freedom in this problem, and I ENTHUSIASTICALLY encourage you to put on your critical-thinking hats, be thoughtful, be experimental, be creative, and be collaborative!

_Here's the deliverable to be turned in next week_*: A report (thoughtfully and professionally formatted)*[5] *in which you explore several questions of your own creation. Just as for the final project, you may work in teams of between 1 and 5 members on this problem.*[6]

Driven by the capabilities of this tool, generate one or several research question(s) and explore them. Off the top of my head, I can think of a couple dozen; I bet you can too! Below, I provide a few examples of research questions. You are welcome to use any of these, but also note that there's so much more fertile ground beyond these examples:

- *How does the timing of the illness peak (the moment of most illnesses) depend on each of the simulation parameters? Pick a certain "neighborhood" of parameters (i.e. set a specific value for each parameter) and explore the sensitivity of the peak to small changes in each parameter by computing a numerical derivative of the peak timing with regard to the input parameters.*
- *Which parameter(s) exert(s) the most influence on the timing of the peak, in the vicinity of your chosen initial values? Note that you may have to average your results over several independent simulations to obtain convincing results.*
- *Even if we reduce our level of social contact, does it also matter whether we are regularly changing our network of social contacts?*

Regardless of your motivating question(s), before diving into simulations, be sure to state your hypotheses. Your final submission should include a mixture of plots, calculations, and exposition.

You are not limited to pursuing questions with an epidemiological flavor! You may also spend time making (and write your report documenting) improvements to this simulation tool (e.g. increasing its efficiency or expanding its functionality). For example, if you are working as part of a team, one team member might develop new functionalities and pass those on to the rest of their team to answer even more challenging questions.

Finally, I encourage you to think of creative ways to connect your work with previous ideas from 12-271 (e.g. numerical calculus, probability and statistics, least-squares fitting). As one example, you might explore whether you can use this simulation tool to predict exponential growth in number of sicknesses, as we saw in the US and Italy COVID-19 data from the previous problem set. When does the growth fall off the exponential curve? What is the reason for this behavior?

---

[5] Professional writing is an area of emphasis in our department, and all students receive extensive instruction in a variety of writing formats before taking this course.

[6] Students were allowed to self-select teams. Of the three "large" teams (four students or more) that formed for this assignment, all three featured diversity in gender and two featured diversity in race, roughly reflecting the level of gender and race diversity in our department.

NB: To set a sense of expectations, you should plan to spend ~4-6 hours of earnest effort per person on formulating questions, developing hypotheses, coding, conducting numerical experiments, writing up your findings, and editing your report. (This recommended window of time of course also includes any time spent in lab working on this problem.) I will not tell you how many independent problems you need to investigate in order to receive full credit; in principle, a single immersive question could easily occupy 30 person-hours of serious effort. Be honest about time spent and be adventurous (from the confines of your room)!

Appendix B: Prompt for "Graph Theory and Gratitude"

As we finish a semester of seriously challenging material – and amidst seriously challenging circumstances – I want to offer all of you an opportunity to submit "Participation Shoutouts!" to express your feelings of appreciation for those other members of 12-271 whose engagement in this class (inside or outside the classroom) has helped you grow as a computational scientist and engineer. This continues (and, in fact, caps off) a semester-long tradition of acknowledging others for their contributions to our learning.

A "Participation Shoutout!" is a statement (positive, thoughtfully written and formatted, detailed, preferably featuring anecdotes, minimum 75 words in length) specific to the contributions of a single person. These are to be submitted via our course webpage.

With your submitted shoutouts, please indicate whether you are comfortable with the material being shared with that individual. If permission is given, all sharing will be strictly anonymous.

To be absolutely clear:

- *You are **NOT** required to submit any "Participation Shoutouts!"*
- *Your participation grade will **NOT** be negatively impacted whatsoever if you do not submit any shoutouts or if you are not mentioned in any shoutouts.*
- *Your participation grade **MAY** increase (by up to 0.5 /10 points) if there is a consistent pattern that peers found your engagement enriching for their 12-271 experiences.*
- *There is no limit to the number of "Participation Shoutouts!" you may submit.*
- *Study after study after study shows that life happiness correlates EXTREMELY strongly with feelings of gratitude. There is literally no downside (and a whole lot of upside, for you and for others) to submitting shoutouts.*