

Network complexity analysis

Erik Amézquita

12/05/2023

- **Mentor:** Erik Amézquita
- **Affiliation:** Plant Sciences & Technology and Mathematics @ MU
- **Area:** Applied Mathematics
- **Start time:** First week of Spring 2024 semester
- **Commitment:** Attend one-on-one regular weekly/biweekly meetings with the mentor, plus a weekly training meeting with BIPS fellows.
- **Duration:** The whole semester with possibilities to extend further depending on progress
- **Compensation:** Compensation through BIPS Broader Impacts Grant. Up to \$12/hr — expected 10-15 hrs a week

Overview

Plenty of real-world data comes encoded as networks: collections of vertices and edges. For instance, the edges could be individual genes and edges indicate certain interaction between them. More often than not, complex traits such as plant height or flowering time are controlled not by one but by multiple genes and multiple interactions. Providing an adequate mathematical description of such a vast array of vertices and edges can be daunting. How “connected” is the graph? How “vital” is an individual node? If we could add a new edge, where would it be the most “impactful”? Are there “typical” motifs repeated over and over again? How do we mathematically define “connected”, “vital”, “impactful”, and “typical” in the first place?

The general aim is to understand some concepts of network analysis and to then apply them to network data from various plant biology-inspired contexts. I am particularly interested in reading the work of [De La Cruz Cabrera *et al* \(2022\)](#). In this paper, they provide ways to analyze contributions of individual nodes and edges to the overall structure of the graph. To the best of my knowledge, these specific network analysis techniques have not been applied yet to any plant biology context. The ultimate goal would be to code —preferably in python— the proposed analysis techniques by De La Cruz Cabrera *et al* and apply them to both social and yeast network data. The whole research project will be **application driven** rather, providing novel applications of math to biology. The emphasis of the project will **not** be on developing new mathematical theory itself.

Although I have listed [De La Cruz Cabrera *et al* \(2022\)](#) as my main motivation paper, we can change the main reading material based on student’s strengths.

This would be a highly interdisciplinary research project, combining mathematics, programming, and biology. Once we have a basic grasp of the complex network analysis side of things, we will be involved in active discussion with plant biologists to truly understand the significance of our results and fuel future directions.

BIPS Funding

The research will be funded by the BioInformatics in Plant Sciences (BIPS) program. BIPS is an interdisciplinary undergraduate program where students from engineering and math work in pairs with plant biology students to work on a project that requires the expertise of both fields. From the start it was clear that students working in wet and dry labs needed to learn how to communicate with each other and this requires language fluency opportunities to understand the challenges and opportunities of both fields.

BIPS students work on their projects ~10hrs/week and are required to attend a weekly meeting coordinated by an MU graduate student (BIPS coordinator) who meets with students to discuss their projects, find solutions and alternatives to bottlenecks within their projects, hear about faculty conducting inter-disciplinary research, career opportunities, demos of new software for data mining, phenotyping, genomics and to prepare posters and talks for presentations.

BIPS students will be compensated up to \$12/hr. US citizenship is **not** required. Funding can extend beyond spring semester depending on progress. BIPS is currently administered by [Dr. David Mendoza-Cozatl](#) from the Division of Plant Sciences & Technology.

Profile of student

The student should feel comfortable with some linear algebra and calculus at the bare minimum. Some knowledge on discrete mathematics or real analysis would be helpful albeit not required. Basic coding skills (e.g. feel comfortable using `numpy` in python) would be very helpful but not required.

The exact path of the research project will vary depending on the student's strengths.

The student should be open to discuss their knowledge with biologists and be receptive of biology input when it comes to data analysis and interpretation.

References

- O. De la Cruz Cabrera, J. Jin, S. Noschese, L. Reichel (2022) Communication in complex networks. *Applied Numerical Mathematics* 172: 186–205. DOI: [10.1016/j.apnum.2021.10.005](#)
- E. Estrada (2011) *The Structure of Complex Networks: Theory and Applications*. Oxford University Press.
- M. Newman (2010) *Networks: An Introduction*. Oxford University Press.