Exploring Effect of Different Network Topologies on Fixation Time in Evolutionary Graph Theory
07-300

Yitao Chen (Viola)
https://violachenyt.github.io/07300.html

November 4, 2021

1 Background and Project overview

I will be working with Professor Oana Carja and her graduate student, Y.P. Kuo, in the Computational Biology Department on exploring the effect of different graph topologies in shaping the time to fixation in an evolutionary setting.

Genetic algorithms and evolution-based optimizations have been gaining increasing attention in recent years due to their huge potential in machine learning applications. Previous studies have been mostly focusing on the probability of fixation, which is the probability for a beneficial mutation to eventually stay in the population. Here, we are interested in studying fixation time - the time for a beneficial mutation to propagate through the entire population, ideally without compromising the fixation probability. Shorter fixation time is important for accelerating rate of discovery in directed evolution and evolutionary optimization.

Previously, Y.P.'s work has studied higher order interactions that involve 3 nodes in the graph [2]. One direction for this project is to explore further beyond interactions of order of 3 and to look into more complex network topologies, and at the same time continue to assess their performance in different fitness terrains. Specifically, we want to examine the performance of the network topologies at circumventing local optimum, since a good evolution algorithm should be able to navigate through local optimum and move towards the global one.

In this project, a Moran birth-death model with fixed population size will be used to track the time and frequencies of mutation in the population. Each node in the graph would represent an individual in the population, and each edge denotes a potential evolutionary corridor, ie. a replacement can only take place if there is an edge between the two nodes involved.

In addition to looking at different network topologies, another direction for the project would be to look at whether different distributions of degrees in the network have effects on the fixation time and fixation probability, since distributions could have the same mean and variance but still be very different.

There are a few potential challenges in the project. Specifically, it might be challenging to formally define network topologies when the order is high and there could be potential challenges with graph isomorphism. Ultimately, most of the challenges perhaps lie in figuring out the set of sensible simplifying assumptions to make and coming up with network topologies to study.

2 Project Goals

2.1 75% Project Goal

- Recreate results from [2] regarding effect of triangles wedge on time to fixation in networks of small mean degree and k-regular graphs. This will help to confirm previous results and ensure the setups are consistent.
- Study the effect of a heavy-tail distribution of degree on fixation time and probability on the existing network topology of wedges and triangles, across various fitness spaces and complex terrain.

2.2 100% Project Goal

- Recreate results from [2] regarding effect of triangles wedge on time to fixation in networks of small mean degree and k-regular graphs. This will help to confirm previous results and ensure the setups are consistent.
- Study the effect of a heavy-tail distribution of degree on fixation time and probability on the existing network topology of wedges and triangles, across various fitness spaces and complex terrain.
- Develop theoretical framework and construct models for studying higher order interactions among 4 nodes, which involves a lot more complicated topologies in the network motif as compared to the case of 3 nodes.
- Construct simulations to examine the fixation time of different network topologies involving 4 nodes across various fitness spaces.

2.3 125% Project Goal

- Recreate results from [2] regarding effect of triangles wedge on time to fixation in networks of small mean degree and k-regular graphs. This will help to confirm previous results and ensure the setups are consistent.
- Study the effect of a heavy-tail distribution of degree on fixation time and probability on the existing network topology of wedges and triangles, across various fitness spaces and complex terrain.
- Develop theoretical framework and construct models for studying higher order interactions among 4 nodes, which involves a lot more complicated topologies in the network motif as compared to the case of 3 nodes.
- Construct simulations to examine the fixation time of different network topologies involving 4 nodes across various fitness spaces.
- Design more complex fitness terrains to better simulate mutations.
- Explore the performance of the different network topologies on the fitness space and look for any trends or interesting discoveries.

3 Project Milestones

3.1 First Technical Milestone

By the start of next semester, I hope I would have a better understanding of evolutionary graph theory and the recent advances in the field. I will also hopefully have the basic code for graphical simulation prepared so that I can make modifications and add complexity to it easily.

3.2 First Biweekly Milestone - Feb 1

By this biweekly milestone, I hope to be able to reproduce the previous results on network motifs of order 3.

3.3 Second Biweekly Milestone - Feb 15

By this biweekly milestone, I hope to have set up a few classes of fitness spaces that covers a variety of local optimum scenarios that I could test the network motifs on.

3.4 Third Biweekly Milestone - March 1

By this biweekly milestone, I hope to have theorized a few possible network topologies of the 4th order to possible explore.

3.5 Fourth Biweekly Milestone - March 15

By this biweekly milestone, I hope to have run some simulations to examine the fixation time of the various topologies.

3.6 Fifth Biweekly Milestone - April 1

By this biweekly milestone, I hope to have explored a few different distributions on the degree of nodes in the graph, and examine the combination of degree distribution and network topologies.

3.7 Sixth Biweekly Milestone - April 15

By this biweekly milestone, I hope to have wrapped up my project and have come up with a formal written document that summarizes the project which I could refer to later on.

3.8 Seventh Biweekly Milestone - May 1

By this biweekly milestone, I hope to have finished wrapping up the project. If done early, I would reflect on potential future directions of the project which I could potentially work on over the summer.

4 Literature Search

Evolutionary graph theory is an exciting area with a lot of potential to explore. My project will be partially based on the work of Y.P. Kuo, the Ph.D. student I will be working with. [5][2] There are also a lot other discussions that bring insights and interesting ideas to the problem. [4][3][1]

5 Resources needed

The main resources that I would be using will be Python and potentially other programming tools such as Jupyter Notebook. Datasets used will be mostly generated by myself or from other publications. I might be using the Lane Cluster of the Computational Biology Department if needed.

References

[1] Alvarez-Rodriguez, U., Battiston, F., de Arruda, G. F., Moreno, Y., Perc, M., and Latora, V. Evolutionary dynamics of higher-order interactions in social networks. *Nature Human Behaviour 5*, 5 (2021), 586–595.

[2] Carja, O., and Kuo, Y. Evolutionary graph theory beyond pairwise interactions: higher-order network motifs shape times to fixation in structured populations.

[3] Gupta, A., Savarese, S., Ganguli, S., and Fei-Fei, L. Embodied intelligence via learning and evolution. *arXiv preprint arXiv:2102.02202* (2021).

[4] Jaderberg, M., Czarnecki, W. M., Dunning, I., Marris, L., Lever, G., Castaneda, A. G., Beattie, C., Rabinowitz, N. C., Morcos, A. S., Ruderman, A., et al. Human-level performance in 3d multiplayer games with population-based reinforcement learning. *Science 364*, 6443 (2019), 859–865.

[5] Kuo, Y. P., Nombela-Arrieta, C., and Carja, O. A theory of evolutionary dynamics on any complex spatial structure. *bioRxiv* (2021).