Analyzing Network Structure Through Temporal Motifs

Clark Mollencop
Mentor: Dr. Abhijin Adiga

Network Systems and Advanced Computing Division

Background: What is a temporal motif and why do we care?

- Informally, a motif in a network is a small pattern. In this project, I was concerned specifically with 3-edge, 2- and 3-node motifs in temporal, directed graphs.
- A temporal motif in a directed graph is one where all edges have an associated timestamp (i.e., there was some connection between the two nodes at some time t). These motifs occur within some time window, typically denoted as delta.
- These motifs can be used to detect anomalies and learn about the structure of underlying networks.
- We are using these as a tool in analyzing networks to predict the spread of Tuta absoluta, an invasive species that affects tomato plants.
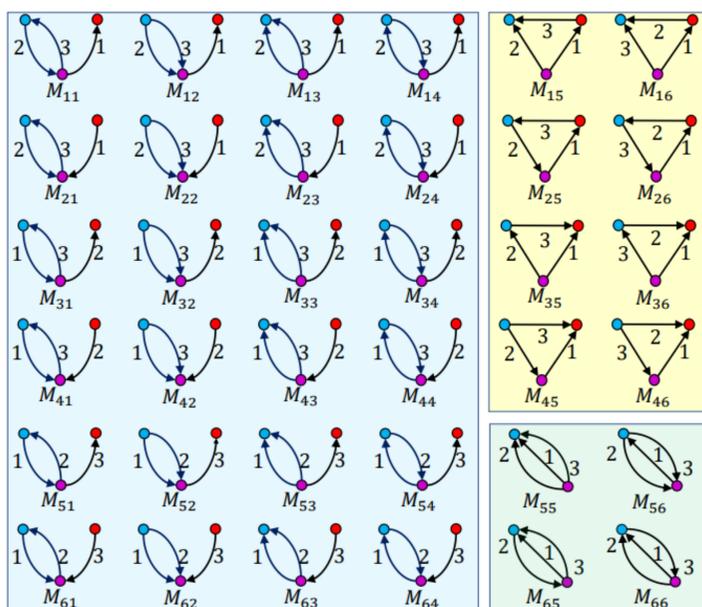
Figure #1: All 2- and 3-node, 3-edge, δ-temporal motifs.

Methodology:

- I built on code written by Gao et al. that finds all of those motifs listed above in linear time with respect to the number of edges. Specifically, I wrote code to figure out when those motifs occurred, as previous work only counted a total number of them in the network.
- I used C++ for the finding of the motifs, and Python (with Pandas and Seaborn for data analysis and plotting) to analyze the results.
- I found all motifs in an actual network of flows of tomato plants in Senegal and predicted flows in Nepal for all time windows 1 through 12 months.
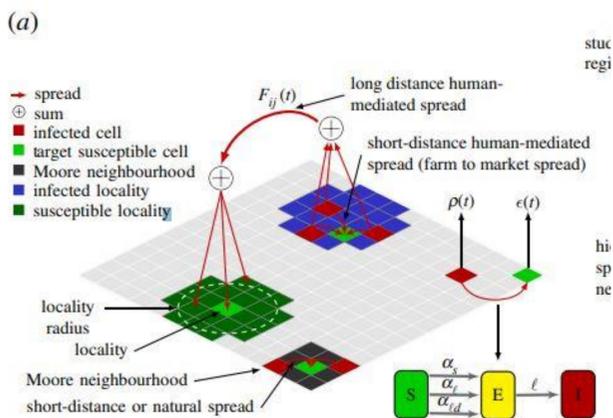

Figure #3: An example of a graph with edge attributes.

Current Work:

- I have added onto work that has created code to find all 3-edge, 2 and 3 node motifs in a directed temporal network. My contribution is finding out when they occurred and analyzing those motifs in periodic networks.
- Periodic networks are those that are expected to repeat after a certain amount of time. For example, growing seasons generally follow a yearly cycle before they repeat again, so a commodity flow network of agricultural trade within a country is similar year to year.
- I also worked on similar code to find 2-node, 2-edge temporal bidirectional motifs.
- The occurrence of certain motifs may suggest periods when networks are especially susceptible to a biological invasion. For example, if there are a lot of outward motifs coming out of a particular node at a specific time, that may suggest that action to stop invasive species spread should be focused there.
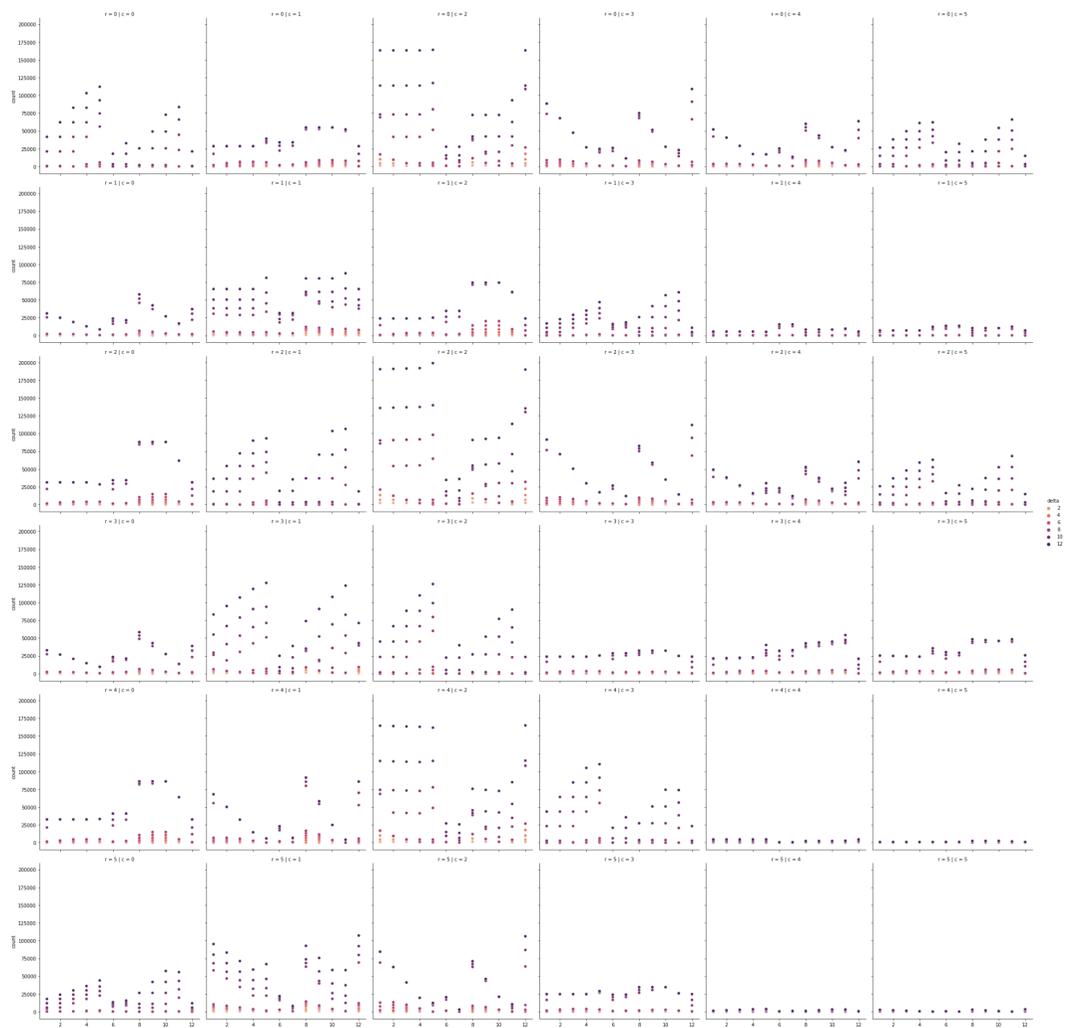

Figure #2: All 2- and 3-node, 3-edge motifs found with deltas 1-12 in predicted network of Senegal tomato crop trade flows.

X-axis: time (months); Y-axis: count; Hue: time window, darker is larger

Future work:

- Expand work to temporal motifs that have edge attributes (e.g., short and long distance spread, human vs. natural spread).
- Larger motifs (4+ edges).
- Defining what exactly a temporal motif means when there are non-unique timestamps.

References

Gao, Zhongqiang and Cheng, Chuanqi and Yu, Yanwei and Cao, Lei and Huang, Chao and Dong, Junyu (2022). Scalable Motif Counting for Large-scale Temporal Graphs. Proceedings of the 2022 International Conference on Data Engineering. https://arxiv.org/abs/2204.09236.

Ashwin Paranjape, Austin R. Benson, and Jure Leskovec. 2017. Motifs in Temporal Networks. In Proceedings of the Tenth ACM International Conference on Web Search and Data Mining (WSDM '17). Association for Computing Machinery, New York, NY, USA, 601–610. https://doi.org/10.1145/3018661.3018731

McNitt J, Chungbaek YY, Mortveit H, Marathe M, Campos MR, Desneux N, Brévault T, Muniappan R, Adiga A. 2019. Assessing the multi-pathway threat from an invasive agricultural pest: Tuta absoluta in Asia. Proc. R. Soc. B 286: 20191159. http://dx.doi.org/10.1098/rspb.2019.1159