**Reaction Paper # 6: Network Evolution**

Mohammad Elkholy  
Computer Science and Engineering  
 The American University in Cairo  
 New Cairo, Cairo, Egypt  
moelkholy@aucegypt.edu

**SUMMARY**

**Empirical Analysis of an Evolving Social Network**

The authors in this paper study the properties of a social network of ~40000 students, faculty and staff at a large university over a 355 day period. They recorded the emails, timestamps, senders, and receivers as well as characteristics of the senders and receivers including a list of classes attended by students and classes taught by professors per semester as well as age, gender, status, departmental affiliation, and the number of years since joining the community. They note that the data collected should not be affected by the personalities of individuals in the network as they will eventually average out. They also filtered out emails that had more than 4 recipients citing it as a way to ensure proper reflection of interpersonal communication rather than mass mailing and spam.

They weigh the edges between two individuals by using the geometric rate of back-and-forth emails in a given time window . They used daily network approximations to calculate the following:

* shortest path length dij
* number fo shared affiliation sij

They also measured cyclic closure (a generalized form of triadic closure) as well focal closure biases (the probability that two individuals will form a new tie given a shared attribute, ex: class). Focal and cyclic closures were only provided for students as we can monitor their attendance.

**Results:**

* Cyclic closure diminishes with dij in absence of shared focus
* Shared focus increases chance of creating ties even if dij > 2.
* mutual acquaintances overcome the absence of shared foci
* likelihood of triadic closure increases with the strength of ties.
* homophily plays a weaker role than expected, gender, status, and time in the community do not have a significant effect on triadic closure
* Smoothing window affects the network measures differently
* Size of largest component, average shortest path length, and average degree change with seasons and different values for , with the average degree being the most affected by
* Clustering coefficient stays relatively constant suggesting that local properties are more preserved compared to global network ones.
* While P(k) stays same across time, the individual degree of nodes changes across time. This suggests that ties, and bridges and their effects change with time.

**Microscopic Evolution of Social Networks**

In this paper the authors try to capture the microscopic properties of networks across time and propose a model that captures these microscopic properties. They use maximum likelihood estimates to match these properties as they had to snapshots of the data in 4 different large scale social networks across periods ranging from 4 months to 4 years. They decompose the network evolution process into 3 different processes:

* Node arrival
* Edge initiation
* Edge destination selection

Their networks are evolved edge by edge and likelihoods of the different processes described above are measured and and their product is the likelihood of the entire model. The higher the likelihood the better explanation it provides for the data.

The authors first consider the bias of node age and degree on edge source and destination selection. They take a look at the preferential attachment (PA) model which suggests that destinations of edges are selected proportional to the destination node’s degree. They compare that with the random Erdos-Renyi network They also study the effect of node age on number of edges it creates by measuring the average number of created edges at each age. They then use MLE to study the combined effect of node age and degree. They use the following MLEs:

* D: probability of selecting a node proportional to its current degree raised to some power
* DR: node is selected preferentially with probability p uniformly at random
* A: probability of selecitng a node depends on its age raised to some power
* DA: probability of selecting a node is proportional to product of degree and age raised to some power.

They find that the preferential attachments model (D) performs best among these across the different networks they studied.

They then study how edges are formed locally and realize that the probability of forming edges h hops away decays exponentially with hop length. This suggests that most local edges created are due to triadic closure. They then study different models for picking which triangles to close, however they find that random selection of both source and destination nodes to close perform the best.

They then study the node lifetime, and find that it is best modeled by an exponential distribution , however the model doesn’t reflect the lifetime of short lived nodes very well. They also study the time gap between edge creation, and find that it follows a power law with an exponential cutoff, the exponential cutoff differs based on the nature of the network. They also study node arrival, and come to the conclusion that there is no one model that can predict all the networks they studied, and thus the model of node arrival should be specified for their model in advance.

Having studied all these parameters, they model their network evolution based on the following:

* Node arrival using N(.)
* node u samples its lifetime with probability
* node u adds first edge to node v proportional to its degree
* node u picks time gap between edge creation based on
* once node wakes up, it lifetime hasn’t expired, it creates a two hop edge using random-random model of triadic closure
* repeat from step 4 if lifetime hasn’t expired

They then provide a proof for their model and test its performance by cutting off a real network G at time T/2, and evolving it using the model they proposed and studied the differences in clustering coefficients, degree and geodesic distance between the real network at time T and the one they evolved and find that they provide a very good fit.

**COMMENTS**

**Empirical Analysis of an Evolving Social Network**

I thought it was very interesting to see how network properties changed, or stayed the same, for different snapshots of the network over time. I did not expect that the clustering coefficient would stay the same, but one thought that comes to mind is that given the academic setting, once someone leaves a group he/she used to be in one class/semester, they join other groups in other classes. The fractional size changes of the largest component makes sense, as different students participate in different clubs, and thus it is very easy to see why the fraction is very big. The fractional size drops a bit for at the beginning of winter break and during summer, which makes sense as many students do leave during breaks to visit family or go on vacations. It was very interesting to see that status, age and gender do not affect the formation of triadic closure as much, although I can not think of a reason for that yet. I would have loved to see a contrast between the changes in network properties versus another dataset of a different setting to see whether the temporal changes are affected by the nature of the dataset or are preserved.

**Microscopic Evolution of Social Networks**

This paper was very detailed. They broke down the network evolution into several processes and started by building their parameters from the ground up based on basic observations to already present social networks. They then provide proof that the model works by taking a network half way through its lifespan and evolving it using their model, then comparing the real network with their model and other models. I would however would have liked to see how the network would have evolved had the they started at t=0. They also changed their definition of node lifetime half way through the paper without explaining much. But other than that, I thought it was a very complete paper.

**REFERENCES**

[1] Kossinets, Gueorgi, and Duncan J. Watts. “Empirical Analysis of an Evolving Social Network.” Science, vol. 311, no. 5757, 2006, pp. 88–90. JSTOR, http://www.jstor.org/stable/3843310. Accessed 21 Oct. 2022.

[2] Jure Leskovec, Lars Backstrom, Ravi Kumar, and Andrew Tomkins. 2008. Microscopic evolution of social networks. In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '08). Association for Computing Machinery, New York, NY, USA, 462–470. https://doi.org/10.1145/1401890.1401948