Critical Review on Empirical analysis of structural properties, macroscopic and microscopic evolution of various Facebook activity networks, written by Khadangi, Bagheri and Zarean in 2017

The main goal of this paper is to develop a close-to-reality activity network model of Facebook. Past research on Facebook social relationships only revealed links between any two nodes without considering whether two people connected on Facebook are close friends with more interactions or just acquaintances. Hence, it is necessary to examine activities and interactions between users to clearly show the quality of such relationships. Users must engage in different activities and exhibit different behaviours with various friends or acquaintances. Then, the authors examined the structural properties of activity networks on Facebook, which included degree distribution, small-world phenomenon, degree correlation, reciprocity, homophily, densification and shrinking diameter over time. They also explored the microscopic and macroscopic evolution of these networks. This paper ultimately offered a comprehensive definition of Facebook's activity network.

In their conclusion, the activity network should be a directed and weighted multilayer network. By examining the structural properties, they found out that the activity network was neither highly reciprocated nor highly assortative. Low reciprocity was due to the asymmetric relationship between any two nodes, which suggested the activity network needs to be analyzed as a directed graph. All the activity must be counted as weighted, since anchoring on different nodes resulted in different measures of structural properties. Furthermore, the small average path length indicated that people have more connections online than offline friends. Thus, researchers must organize the network into weight and multilayer and better understand the specific type of relationships between different nodes. They also found out that degree distributions of all activity networks follow a power-law distribution.

The authors collected data of 36,204 users over three years. I think such a longitudinal study is beneficial in revealing real evolutionary patterns of behaviours and activity networks over a long period.

What I like this paper is its approach to Facebook as a directed network, which differs from previous work. Facebook, unlike Twitter or Instagram, is usually understood as an undirected graph since we can add people to become friends. In contrast, Twitter and Instagram users can follow others, and others do not need to follow back. Fundamentally, friendship is

closer to a directed relationship rather than an undirected relationship. Hence, previous research examining Facebook as an undirected network is inappropriate, which does not reveal relationships in reality. Even though in Facebook's setting that people can either choose to become friends or not friends, as an undirected relationship, we still need to examine such a relationship more carefully. This paper's approach is very bright through studying reciprocity to understand the asymmetric social relationship in the network.

I also agree that network on Facebook must be weighted and multilayer since users such as key-opinion-leader (KOL) and other types of leaders must have more considerable social influence than others. Edges connecting to these leaders must have different weights to make the network close to reality. By having different weights on edges, the network graph should not be a single layer. Thus, a multiplayer network is indeed a better, dynamic and realistic representation.

However, some points need more convincing explanations. In section 4, the authors gave analyses of the structural properties of the activity network. They plotted graphs to check the power-law distribution of various activities, and bar graphs of path length in different activity networks. My first question is why all these activities follow a power-law distribution and what are the insights behind this phenomenon. What we learned from class about power-law distribution arises from people's decisions influenced by others instead of making independent decisions. However, the authors did not provide qualitative explanations of why power-law distribution happened in interactions of social relationships.

The analyses are all quantitative, in which authors only compared graphs without giving qualitative explanations of what happened in Facebook users causing such phenomenon. For instance, on page 256 and paragraph 3 describing figure 5, the authors indicated that "most path lengths in mixed and like networks are lower than five, and the average path length in these networks is very low." More qualitative explanations, elaborating on what might happen among users, can convince readers that the approach is appropriate. I propose that users have more interactions with cyber acquaintances rather than solely interacting with close friends. Therefore, by having more acquaintance as having more weak edges, the 6-degree separation will shrink due to a larger span of connections with more people.

In section 5, the authors provided longitudinal analyses on those structural properties. They studied the shrinking diameter, densification, clustering coefficient, reciprocity and homophily separately. However, it will be more convincing in considering all these properties in

a holistic way. I think there exists a strong relationship between densification, shrinking diameter property and clustering coefficient. In section 5.2, the authors mentioned that densification rapidly increased for “like” activity, and the average distance between nodes shrank over time. In section 5.3, the clustering coefficient of the four activity networks all increased over time. The increasing clustering coefficient may entail that more people got to know and interacted with each other, which would explain the increase in densification. This could further cause the average distance between people to decrease, so the way to get to a target person became short. It is essential to integrate all the properties to examine the dynamic relationship to gain more insights instead of separately study each property.

In terms of limitations, this research collected 36,204 users, which might not be representative. In the methodology, the authors did not clarify how they selected users and threads at the beginning. Randomization is critical for data collection. From a statistical perspective, if the age groups of 36204 users are mostly between 20 to 30 years old, then the data is insignificant in representing all social relationships. I suggest the authors must clarify how they randomize their data collection to accurately represent and cover a more extensive range of different types of age, occupation, genders and ethnicity.

This paper is well-organized. The authors provided the necessary definition of each structural property in the literature review. Before introducing their methodology, the authors summarized others’ approach as a big picture of previous works’ approach, which is very helpful. They examined the network both from structural properties and change in properties over time with specific quantitative analysis and graphs. Ultimately, they provided a clear conclusion from a general view to define the activity network and a detailed explanation of each property. They also critically provided limitations on their study for future research.