Evolution of Social-Attribute Networks: Measurements, Modeling, and Implications using Google+

Summary Review

(Dated: February 10, 2016)
**Purpose:** This paper explores the interplay between user attributes and links in a social network. The basic model is the union of a social graph with a bipartite graph capturing user-attribute associations. They have collected an impressive dataset, crawling the full (publicly visible) google+ social network. They extend the usual macroscopic network metrics (degree distribution, density, transitivity, diameter, assortativity) to attribute-labeled graphs, and make several interesting observations about the impact of attributes on network evolution. They then define a model where node attachments are driven both by social connectivity and by attribute proximity. This model is shown to match the main macroscopic features of the data, and to perform slightly better than [55] as a network simulator.

**Dataset:** Google+ network during its first few months of operation. They crawled it in daily fashion. Starting from 6th July 2011 to mid October 2011. They have 79 daily snapshot within 100 days. Node size increases from 1M to 30M at the last snapshot and Edges grows to half Billions of interaction. Due to private circles and attributes, there is some kind of bias in their data. They also captures 4 attributes including city, major, school and employer for each users, which for 22% of users they were publicly available.

**Observations & Results:**

- G+ has hybrid nature, combination of traditional social networks and publisher-subscriber ones (e.g. Twitter).

- Having said that, Network models are not consistent with patterns that exists in 3 phases of G+ network (early days, day 1 to 20, stable growing days, 20 to 75 and open to public)

- Diameter is growing and shrinking and growing.

- Clustering Coefficient decrease, increase and decrease again.

- lognormal is the best fit for in/out degree distribution in stead of power-law. So more low degree nodes than power-law.

- Attribute Nodes have some effects on social structure, such as having more attributes in common can increase the chance of bi-directional link creation, despite number
of neighbors. Or some specific attributes are more effective in creating communities (higher clustering coefficient)

• Number of attributes that are in common between two nodes, has a linear relation with edge creation, in addition to be proportional to degree.

• Considering attribute nodes also help us to predict link creation better in phases after joining the network.