Verifying the Logarithmic Dimension Hypothesis

Anthony Bonato

On-line social networks (OSNs) such as Facebook and Twitter are prevalent in society, and have transformed our paradigms of communication and social interaction. Unlike other complex networks such as the web graph, relatively few models have been posed and rigorously analyzed for OSNs. The recent Geometric Protean or GEO-P model [1] exploits a hypothesized underlying social space, where nodes are close via a prescribed metric if they share similar attributes. In this model, nodes are linked with probability based on both their relative proximity and via a ranking scheme. A key prediction of this model is that a small number of network statistics (including network order, power law exponent, average degree, and diameter) give an estimate on the dimension of the network; that is, the least number of attributes needed to identify agents in the networks. The dimension is predicted to be approximately the logarithm of the number of nodes. This so-called Logarithmic Dimension Hypothesis (LDH) would therefore predict a relatively small number of attributes needed to identify agents, and may have many potential applications to social networks.

To date, there is relatively little evidence for LDH. Independent of [1], a logarithmic dimension of OSNs was also conjectured by [3] for purely theoretical reasons in their model. Tian's Masters thesis [5] gave some preliminary evidence for the LDH using spectral methods.

The main problem that will be considered is finding conclusive evidence for the LDH. A possible approach is to use model selection techniques exploiting machine learning. Using the approach of [2] and [4], the GEO-P model will be contrasted with sampled OSN network data using a number of network classifiers. For control, other networks models such as the binomial random graph model, random geometric graphs, and preferential attachment model will be studied using the classifiers.

References

[1] A. Bonato, J. Janssen, and P. Prałat. Geometric protean graphs. *Accepted to Internet Mathematics*, 2012.

[2] J. Hurshman, M. Janssen and N Kalyaniwalla. Model selection for social networks using graphlets. *Technical Report Dalhousie U, CS-2011-07.*

[3] M. Kim and J. Leskovec. Multiplicative attribute graph model of real-world networks. *Algorithms and Models for the Web-Graph*, pages 62–73, 2010.

[4] V. Memisevic, T. Milenkovic, and N. Pržulj. An integrative approach to modeling biological networks. *Journal of Integrative Bioinformatics*, 7(3):120, 2010.

[5] A. Tian. Models and mining of on-line social networks. *Masters Thesis, Ryerson University*, 2011.