**Project Proposal: Finding “Seed” Users with Maximum Influence in Their Social Circles**

**Group Members:**

Mark Weaver

Maxwell Yi

**Problem:**

Consider the problem of a new marketer looking to exploit an existing social network, to identify which users would be the most effective “seed” users to maximize the number of users aware of a product by propagating that information to targeted social circles and groups. To do this we need to define a machine learning task of identifying users’ social circles. We pose this problem as a node clustering and optimization problem on a user’s network, (a network of connections between their friends.)

**Previous Work:**

Previously, community detection has been performed with the use of classical algorithms that tend to identify communities based on node features or graph structure, but have rarely used both to uncover mixed-memberships to multiple groups. We propose a model that performs clustering on social-network data to model memberships to multiple communities or (social circles) to pinpoint influential users.

**Project Milestones:**

·      Acquiring Data sets

·      Data classification

·      Data parsing: sorting of data

·      Building of machine learning algorithm: finishing models

·      Gather Results

**End Goal:**

The outcome of this project will be evaluated by how accurately a model can identify which users would be the most effective “seed” users to maximize awareness of a product based on the users’ quantity of social circles and their content, determined by an unsupervised method that learns which dimensions of profile similarity lead to densely linked social circles to automatically determine both the number of social circles as well as the social circles themselves.

**Reference:**

*Barbieri, Nicola, Bonchi, Francesco, and Manco, Giuseppe. Topic-aware social influence propagation models. Knowledge and information systems, 37(3), pages 555–584, 2013.*

*D. Hallac, J. Leskovec, S. Boyd. Network Lasso: Clustering and Optimization in Large Graphs. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2015.*

*David Kempe, Jon Kleinberg, and Eva Tardos. Maximizing the spread of influence through a social ´ network. In KDD, pages 137–146. ACM, 2003.*

*Goyal, Amit, Bonchi, Francesco, and Lakshmanan, Laks VS. A data-based approach to social influence maximization. Proceedings of the VLDB Endowment, 5(1), pages 73–84, 2011a.*

*J. McAuley and J. Leskovec.* [*Learning to Discover Social Circles in Ego Networks*](http://i.stanford.edu/~julian/pdfs/nips2012.pdf)*. ACM Transactions on Knowledge Discovery from Data (TKDD), 2014.*