

## **Advanced Lab 2: Implications of Network Structure of the SAP Online Knowledge Community Platform**

In this optional advanced lab, you will examine a network drawn from the SAP Online Knowledge Community platform (<http://go.sap.com/community.html>). This is an induced sub-network based on 10% of nodes drawn from the dataset generously provided by Prof. Peng Huang, of the University of Maryland. See here for more on his research studies: <http://penghuang.com/research/>

The provided CSV file is a directed edge list. The right-hand column represents the ID of a user who posts a question that starts a new thread in a SAP community user forum. The left-hand column represents the ID of a user who provides an answer to the posted question. So, a directed edge from the left-hand node to a right-hand node represents an answer provided to a question. Since a user can answer multiple questions posted by another user on one or more threads, the file contains duplicate edge-pairs, which can be combined to form directed edges of varying weights (using the simplify function in R).

Please submit a report of 3 or fewer pages, that describes the features of the network, its overall structure, the distributions of network measures among nodes and edges, and most importantly, what does the network structure suggest to you about the SAP knowledge community? Consider the general tension between network cohesion and diversity, and what the relevant metrics might mean for trust, reciprocity, authority, and access to information. Also, consider to what extent strong triadic closure is present in this network, and what the data tells us about the role of local bridges and weak ties. For more background, refer to the paper by Aral and Van Alstyne (2011), posted with this assignment, as well as the other optional session readings by Burt (2004) and Granovetter (1973).

Please build on the R script file provided to create any additional diagrams, plots, and statistical analysis used in your report; and submit your final R script (with inserted comments) along with your report.

### **Background readings**

Aral, Sinan, and Marshall Van Alstyne. "The diversity-bandwidth trade-off." *American Journal of Sociology* 117.1 (2011): 90-171.

Burt, Ronald S. "Structural holes and good ideas." *American journal of sociology* 110.2 (2004): 349-399.

Granovetter, Mark S. "The strength of weak ties." *American journal of sociology* (1973): 1360-1380.

Huang, P., A. Tafti, and S. Mithas. "Sponsor Investments and User Contributions in Knowledge Communities: The Role of Knowledge Seeding," *MIS Quarterly* (42:1) 2018, 213-240.