Assessment of Kossinets and Watts (2009)

The main research question of the paper Kossinets and Watts (2009) is: what is the origin of Homophily in an evolving social network? To be more specific, the environment of this research is inside a particular university community. There are two theoretically transmission mechanisms as depicted by former literatures, "individualistic" and "structuralist". The authors would like to test if one single mechanism of these two could account for the homophily observed in the university, or if they separately play partial role and are mutually complementary.

The dataset utilized by the authors was constructed by merging three different databases. There were 30396 individuals and 7,156,162 exchanging messages included in this dataset. Time covers one specific academic year with overall 270 observation days. The definition of all the variables could be found in Appendix A and some of the descriptive statistics are reported in Appendix C and Table1 and Table2 in the body. The variables could be divided into four parts, which are personal characteristics, organizational affiliations, course-related variables and e-mail-related variables.

For the data cleaning process of this article, it is mentioned by the authors that the final dataset included only messages that were sent to a single recipient. This may trigger negative bias of interactions among specific pairs of individuals in the sample. Although the authors stated in the footnote that they repeated the analysis including e-mail with up to five recipients as interpersonal communication, the negative bias problem still exists to some extent.

In this paper, there is a seeming mismatch of the underlying theoretical construct, "social relationships", and the data source, e-mail logs linked to other characteristics of the senders and receivers. The problem here is that social relationship is persistent,

continuous and evolving in time, but the E-mail exchanges is discrete and intermittent. The authors took advantage of an effective method, sliding window filter, to deal with the problem. The essential part of this method is to calculate the geometric average of the number of messages exchanged by users i and j per unit of time, summed over the past t time units, which serves as a conservative measure of intensity. (Kossinets and Watts 2009, p413)

Reference

Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. American journal of sociology, 115(2), 405-450.