

3. Assessment of Kossinets and Watts (2009)

- (a) Why do people tend to group up with those of the same race?
- (b) The data used in this research is collected from a large university in the U.S., including 30,396 undergraduate and graduate students, faculty, and staff. Three different databases comprise the data set: (1) the logs of email interaction within the university, (2) individual attributes such as status, gender, age, etc. (3) records of course registration. The time period span for one academic year. The description and definition of all the variables are listed in the appendix A.
- (c) Since the error-correction strategies are based on longitudinal nature of data set, the potential problem for data cleansing process used in the research might wrongly interpolate some data which might not necessarily relate to time. For example, as the paper mentions that we could use backward interpolation for campus, dormitory, school, etc. However, those attributes are not necessarily related to time. For example, individual's preference might have greater influence on their choice. Therefore, the researchers could get a wrong or biased speculation due to wrongly interpolating the data, which could weaken the hypotheses proposed by the authors.
- (d) One weakness of using e-mail logs to observe social relationship is that this data is incomplete. Due to the privacy consideration, the access of the content of emails is limited. Therefore, the authors cannot really analyze the content of those emails and develop a more precise relationship

between emails and the characteristics of senders and receivers. The solution proposed by the authors in order to address the problem is informed consent procedure. Through exchanging some benefits, some users might be willing to provide their content and data for the researchers. Then, the researchers can better interpret the relations between those e-mail contents and the characteristics of the individuals.