Examining author endogamy as a measure of conference and journal quality
Utkarsh Kasana
Advisor: Dr. Carlos Rivero

Problem:
Examine the role of endogamy in the scientific community, by determining the quality of any journal or conference based on the endogamy among authors.

Endogamy
In the classical sense, it refers to the custom of marrying in one’s own community. In the scientific community, endogamy refers to the degree of collaboration between different scientists or researchers in proposing a new idea or research.

Endogamy Calculation:
- Get all combinations of the author set of a paper.
- Endogamy of each author set:
  \[ E(A_i) = \frac{\sum \text{Common Papers}}{\sum \text{Total Papers}} \]
- Paper endogamy: Average endogamy of all author sets.
- Conference endogamy: Average endogamy of all papers.

Dataset
DBLP Data Repository. [http://dblp.uni-trier.de/](http://dblp.uni-trier.de/) 3 million publications, 1.7 million authors, 4,800 conferences and 1,500 journals.

Implementation platforms:
- R – Dataframes, vectors and lists.
- Java – Neo4J Graph Database, MySQL and MapDB.

Graph Database coupled with Collections based database provided much faster implementation for our huge dataset

Design and Implementation

Results

Present Analogy:
- Good Endogamy -> Bad Conference
- Bad Endogamy -> Good Conference

The extracted endogamies are compared with rankings provided in ERA (Excellence for Research in Australia).

- Top 9 Conferences with lowest endogamies.
  - Positive results and follows the present analogy.
  - Most conferences are top tier.

- Top 9 Conferences with highest endogamies.
  - Does not follow the existing analogy.
  - Most conference are still top tier.

Future work
Even though low endogamy often points to good conferences, it is not always the case. There are other factors like Enablers and New Researchers which we need to explore before determining the quality of any conference or journal.

References
Sergio Lopez Montolio, David Dominguez-Sal, Josep Luis Larriba-Pey, Research Endogamy as an Indicator of Conference Quality, SIGMOD Record, June 2013.