

Co-commenting Network of Restaurants

Jiahui Bi, Xue Cao, Xin Chen, Haochen Liu

Instructor: Feng Mai

Introduction

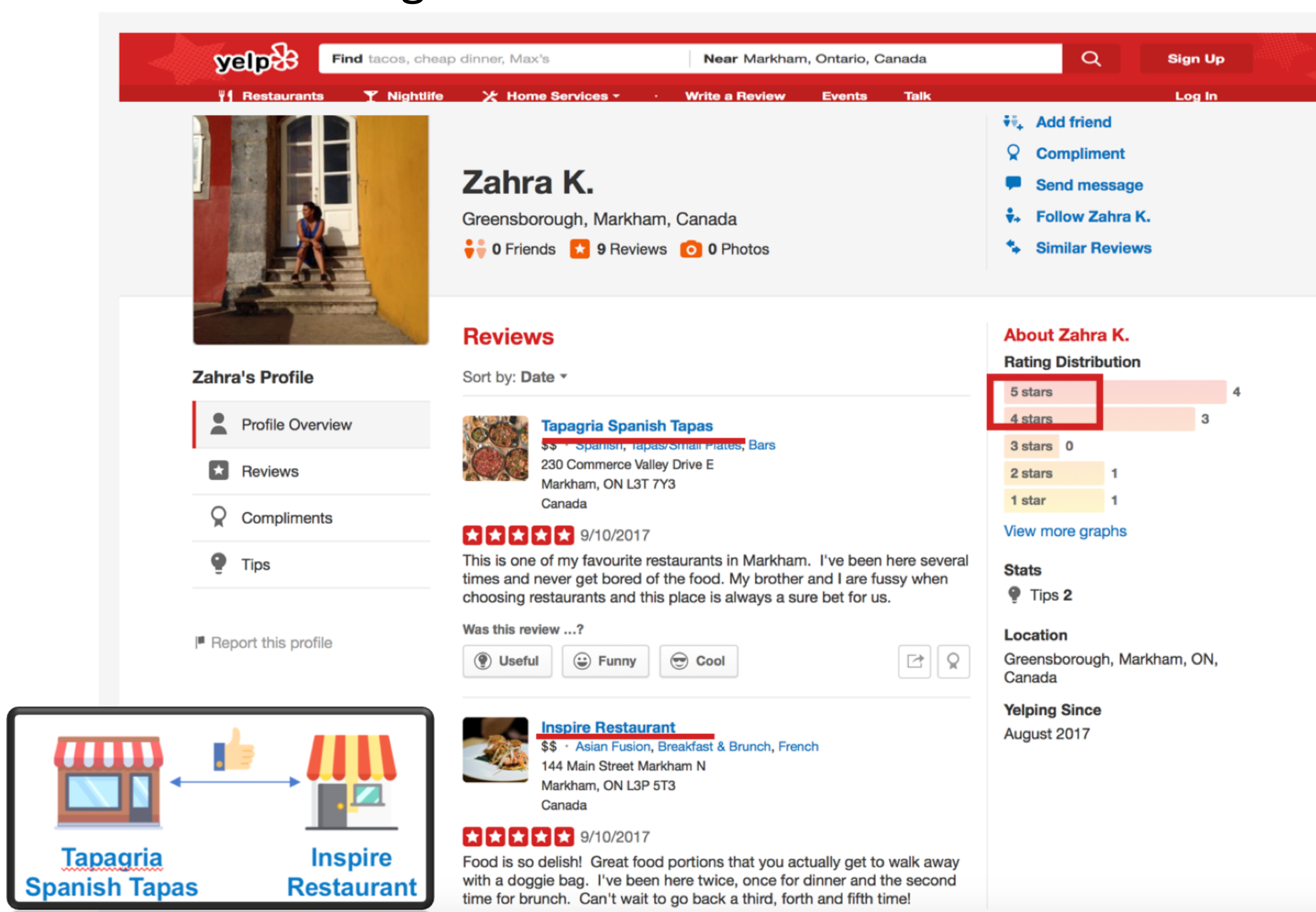
- Although Yelp does not serve as a typical social network supplier, it still can be built as a network which can be formed by business. If one user rated two different restaurants, then we consider these two restaurants have been connected.
- City: Markham

Objective

- Find widely-accepted restaurants
- Analyze local restaurants features

Network building

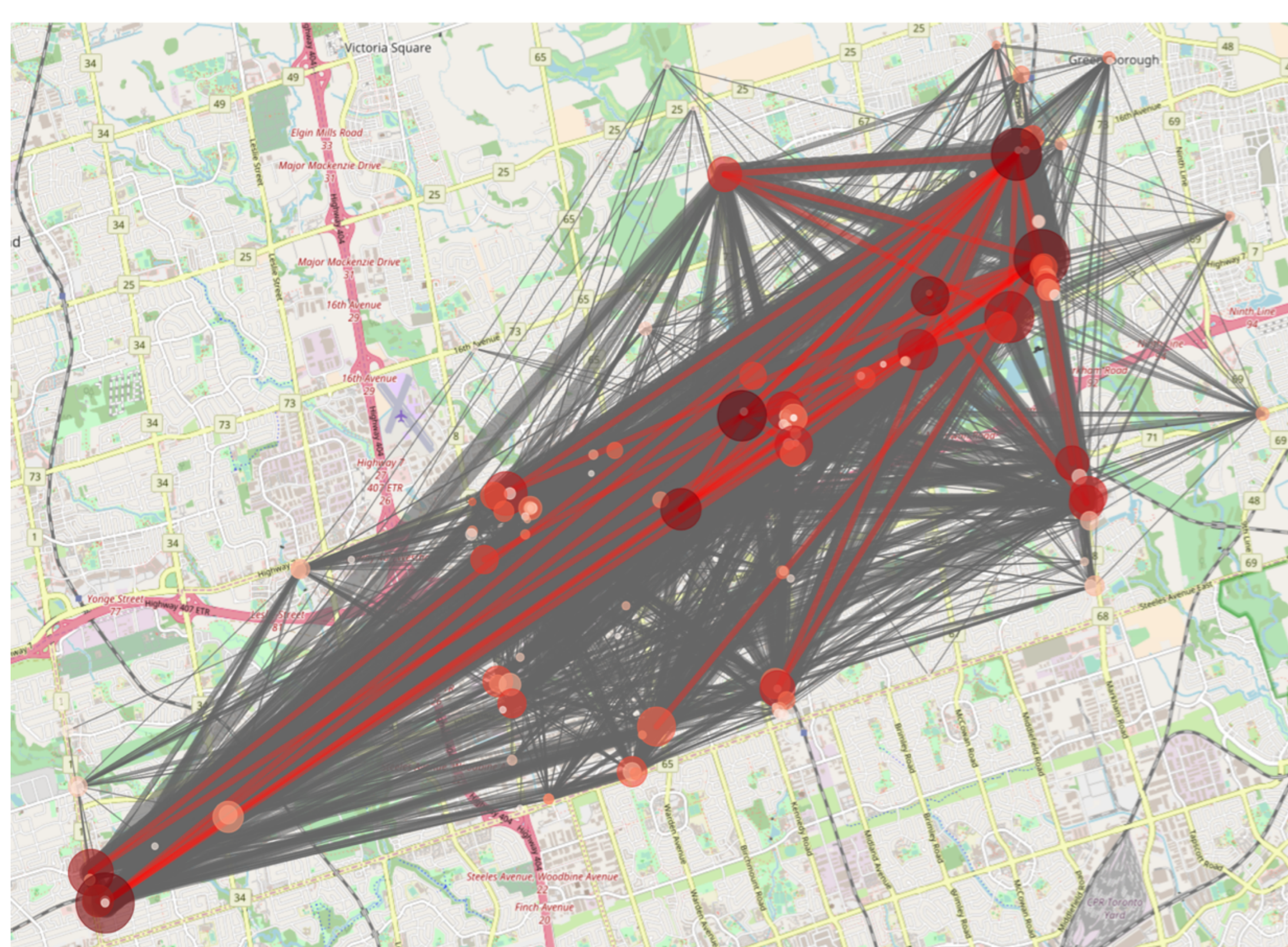
- Co-occurrence network: paired presence
- Co-commenting network



- Restaurants are connected if they got appreciative comments by the same person.



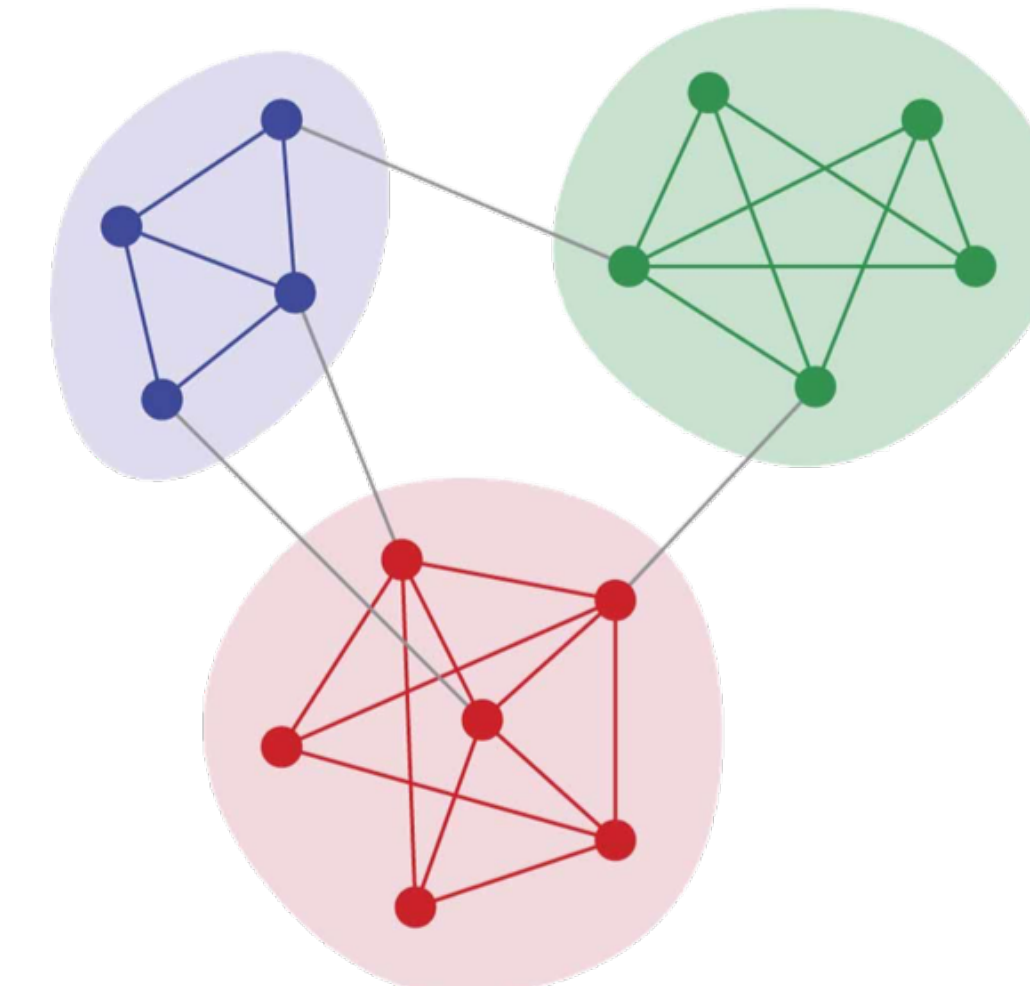
- Highly-connected network and no significant clusters
- Spatial Distribution



- Node color: degree
- Node size: the number of reviews
- Edge width: weight
- Highlight: edge weight is over 50

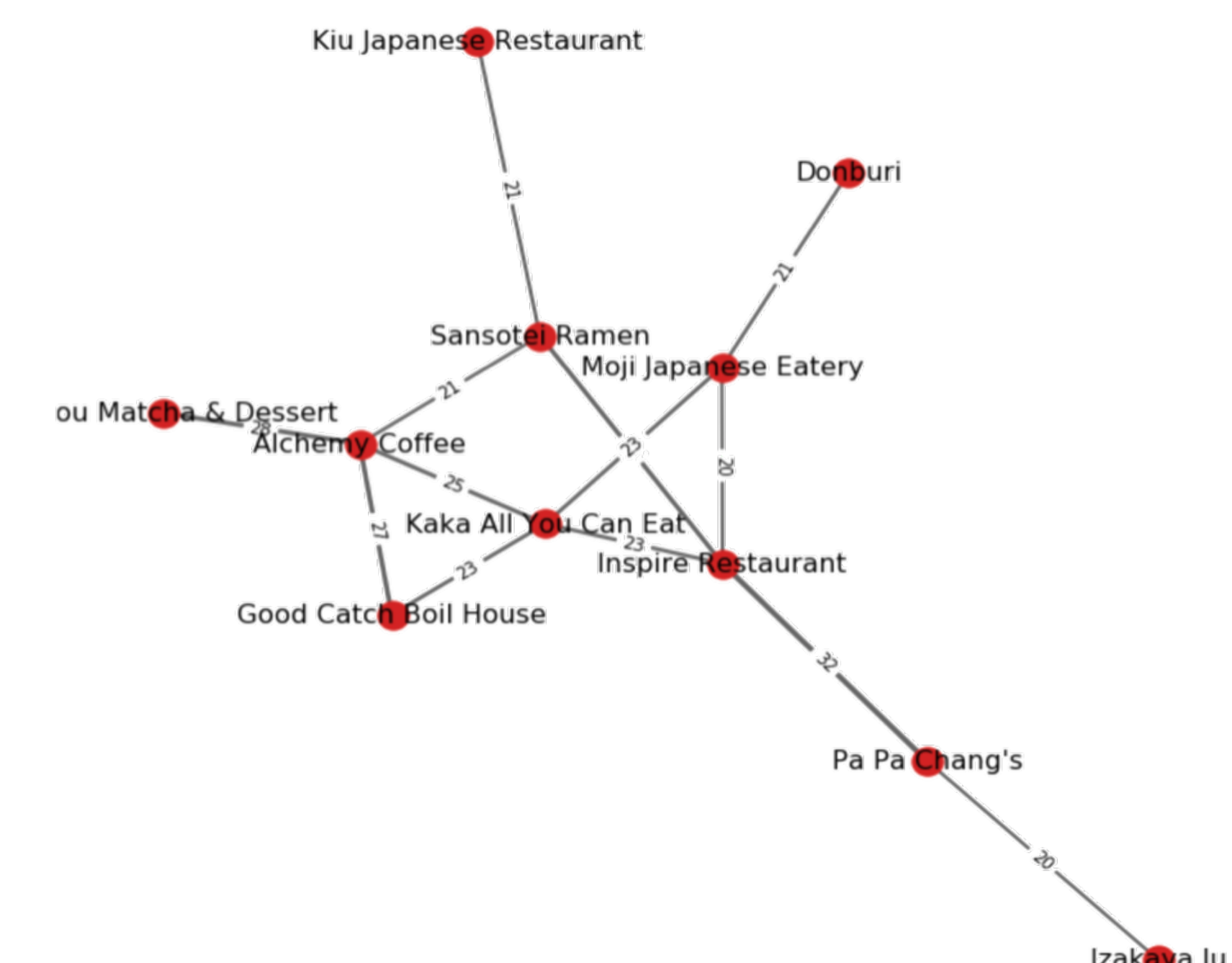
Community Detection

- Detection Algorithm: Girvan-Newman



- Modularity: denser in-group connections and sparser the inter-group connections.
- Maximize the modularity
- Reduce the network size by removing edge's weight below 5 which gets modularity equals to 0.172
- Reduce the network size by removing edge's weight below 20 which gets modularity equals to 0.328
- People dining choices won't present significant clusters.

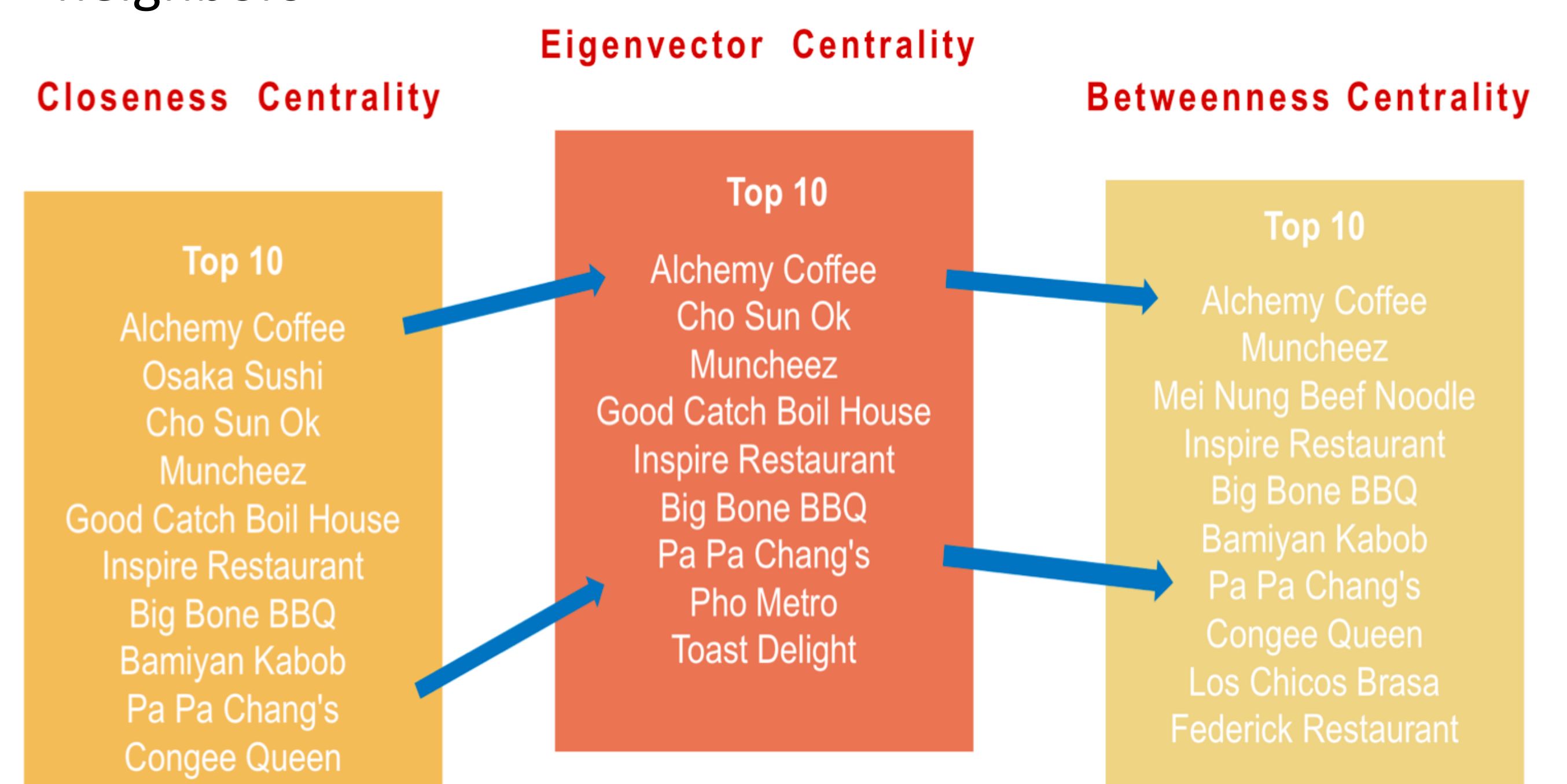
Strong ties



- Asian food
 - around 45% population in Markham is Asian
- Long history

Centrality Ranking

- Closeness Centrality: the shortest paths between all nodes
- Betweenness Centrality: the shortest path between other nodes
- Eigenvector Centrality: the sum of the centralities of its neighbors



- Ranking of different centrality metrics does not change much.
- Some restaurants appear in all three ranking.
- Some above-listed restaurants are also in the yelp official best 10 restaurants.