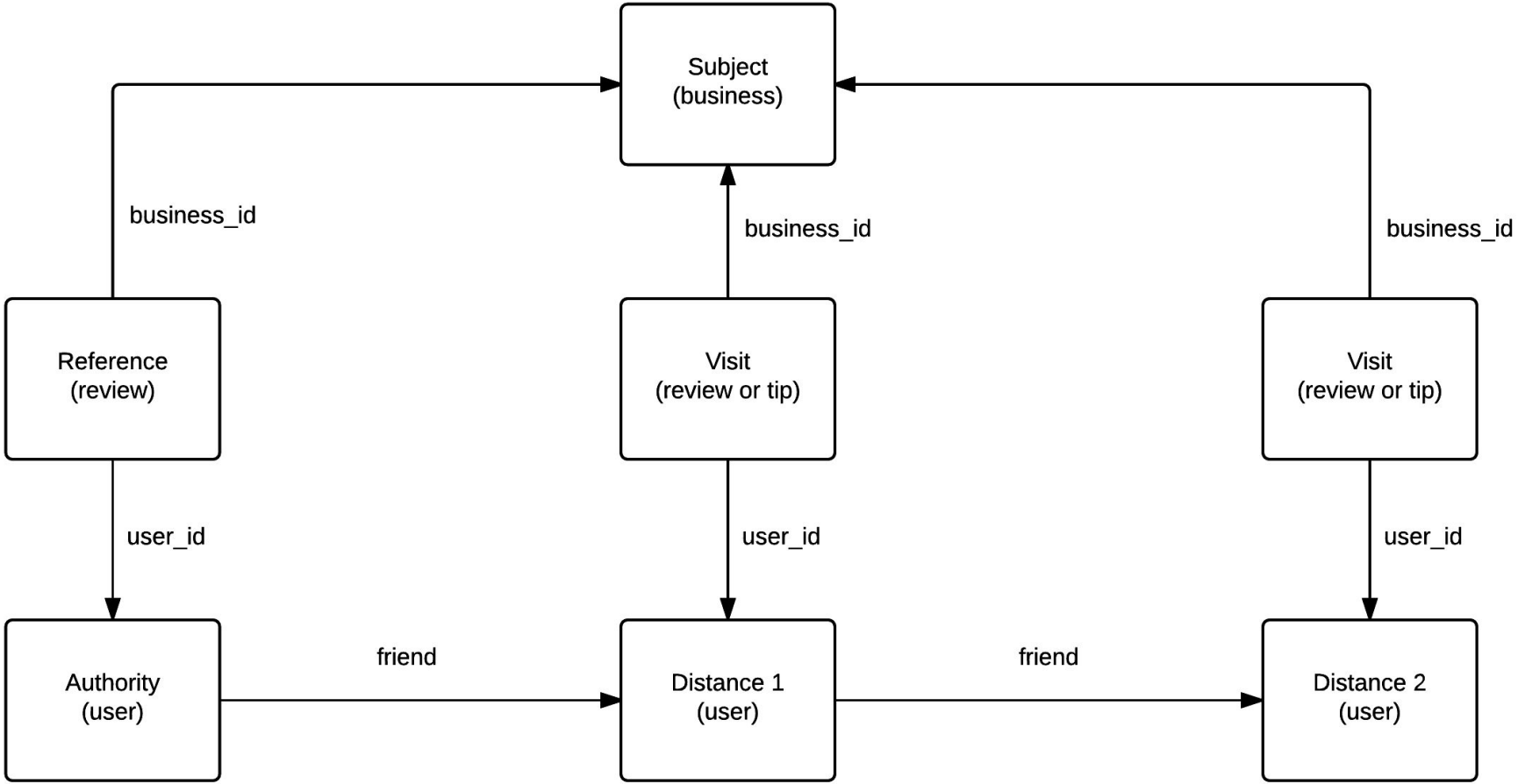
Distant Authorities

Anas Al-Khatib
Jun-Duo Chen
Aymeric Grail

Research topic

- Motivation
  - Information drift over distance
- Subject
  - Distance from authority vs probability of acting on recommendation
- Problem Parameters
  - Yelp dataset
  - Reference posts
  - Distribution of visits


Subject
(business)
business_id
business_id
business_id
Reference
(review)
Visit
(review or tip)
Visit
(review or tip)
user_id
user_id
user_id
Authority
(user)
friend
Distance 1
(user)
friend
Distance 2
(user)

Methods

- Reference posts: classifiers
  - Trainer: “Local experts”
  - Comparison set: “Elite” reviews
- Distribution
  - Effective closeness
  - Actual distance
- Correlation
  - Control set
  - Data aggregation

Progress

- Development environment
  - Drill
  - Python (jupyter, sci-kit, pydrill, numpy)
  - Restricted dataset (Montreal restaurants)
- Reference posts
  - Trainer set
  - Evaluation set