Replication of:
Exploiting Social Network Structure for Person-to-Person Sentiment Analysis

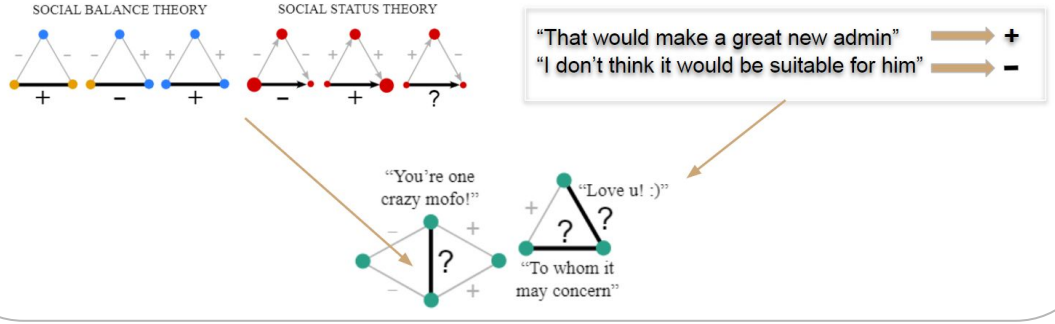
Robert West, Hristo Paskov, Jure Leskovec, and Christopher Potts.

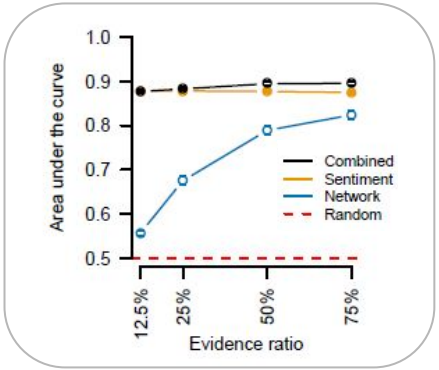
Stanford University

Quick Recap of the Original Paper

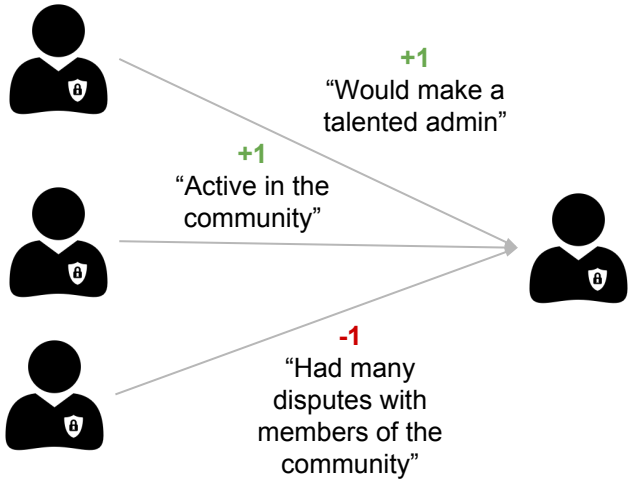
Can Person-to-Person Sentiment modeling be improved using both:

Signed-Network Analysis and **Sentiment Analysis** ?





Details of replication : Wikipedia Dataset



11199 nodes
172062 directed edges

Sampling
of 10 subgraphs, 200 nodes using *BFS*
(2925 edges on average)

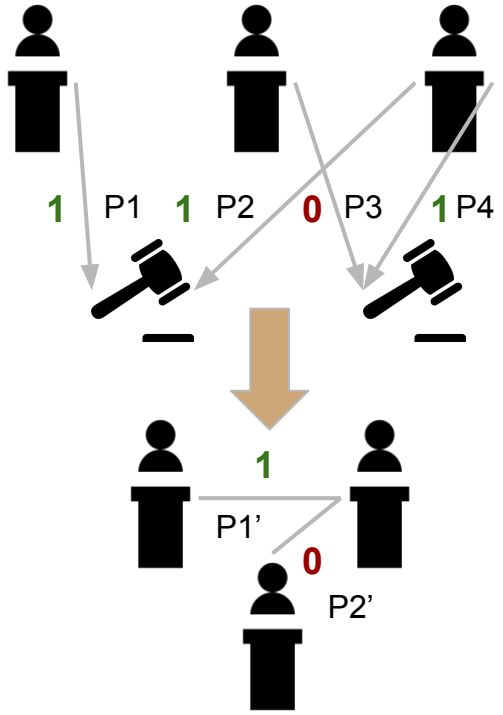
Sentiment Classifier
using *Logistic Regression* on the comments

- Remove stop words, stemming
- Select 10000 most frequent words as features
- Sample 1000 comments with labels for training

Network Analysis
Probability Soft Logic package - CLI

- High weight on prior
- Social Status theory

Details of replication : Convote Dataset



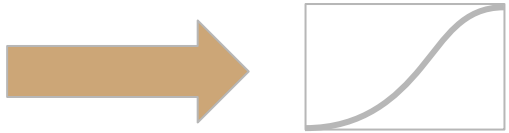
$$P1' = P1 \times P2 + (1 - P1) \times (1 - P2)$$

267 nodes
14690 directed edges

Splitting
5 subgraphs with edges picked *randomly*

Probability distribution transformation
Platt Scaling

SVM output



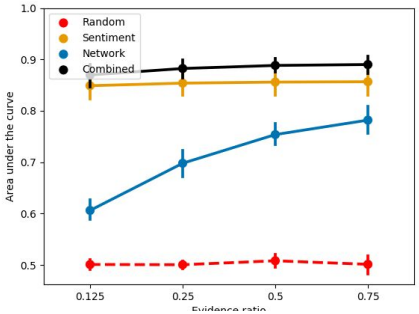
Network Analysis
Probability Soft Logic package - CLI

- Equal weight prior and Network Structure
- Social Balance theory

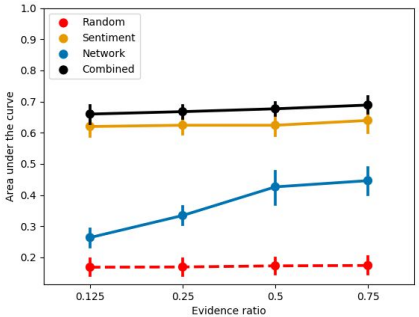
Results of replication : Wikipedia Dataset

Replicated Results

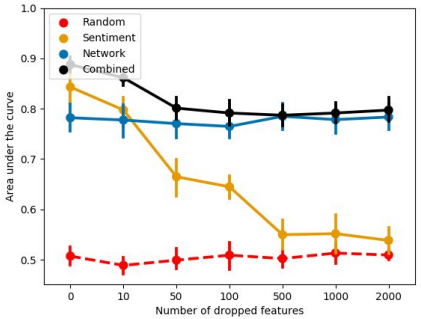
AUC ROC vs Evidence Ratio



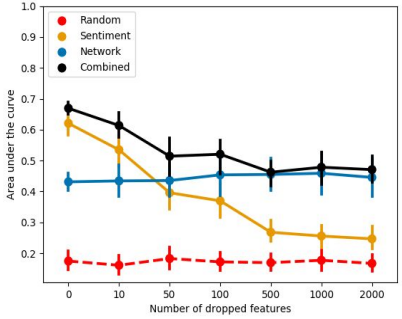
AUC negPR vs Evidence Ratio

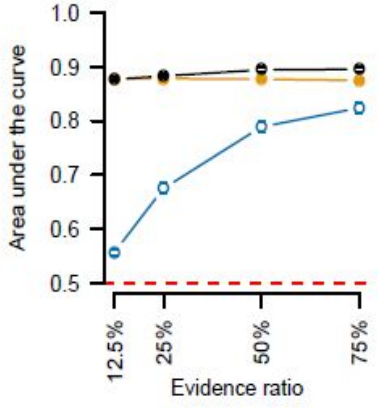


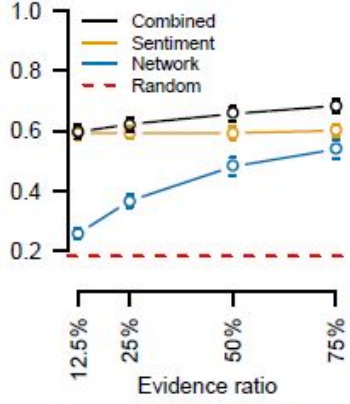
AUC ROC vs # Dropped Features

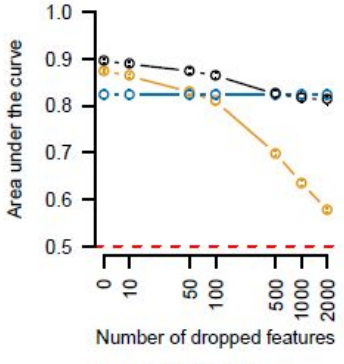


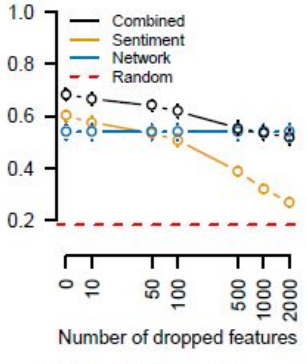
AUC negPR vs # Dropped Features









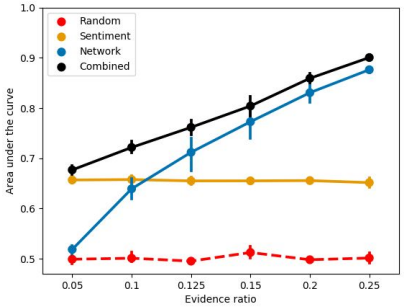


Original Results

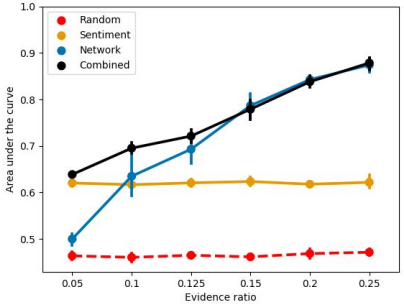
Results of replication : Convote Dataset

Replicated Results

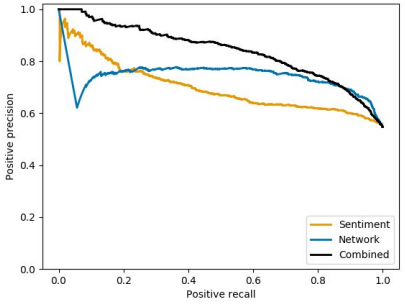
AUC ROC vs Evidence Ratio



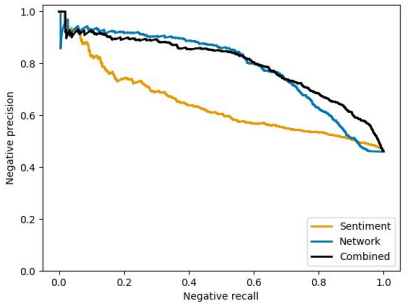
AUC negPR vs Evidence Ratio

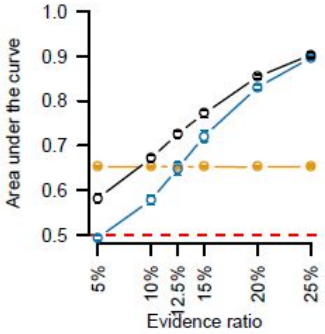


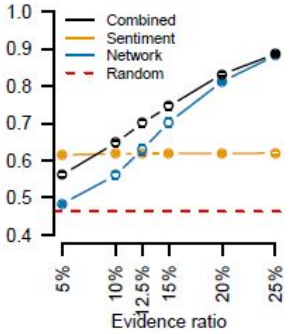
Positive PR curve

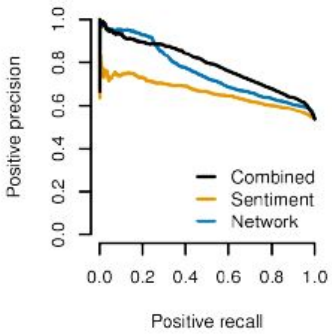


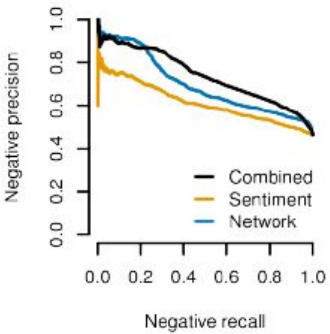
Negative PR curve







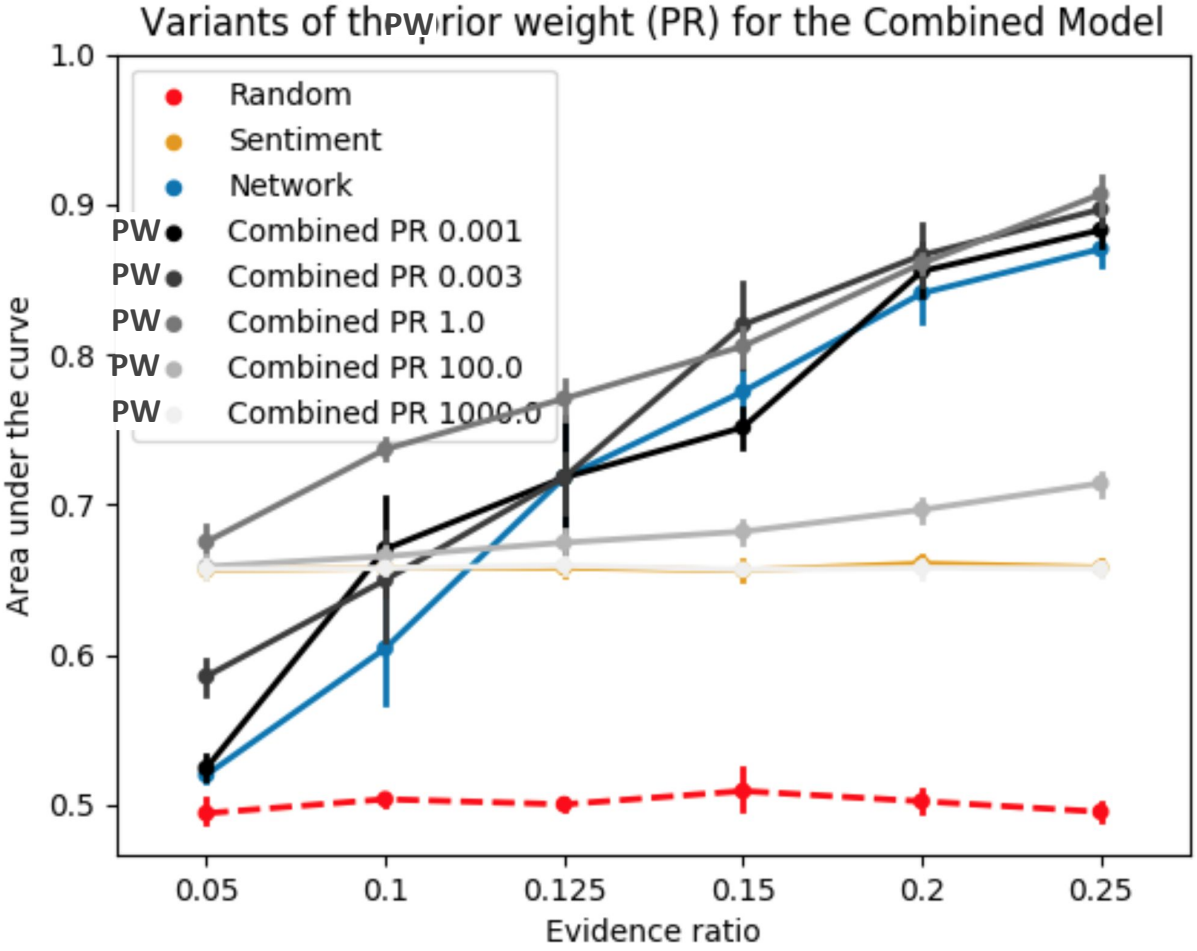




Original Results

Future Extensions

- Learn **automatically** the weights on the Sentiment and Network Structures
- **Non-uniform** weight for the Sentiment probability
- Try other **Network Rules** to discover new patterns


Variants of the prior weight (PR) for the Combined Model
Random
Sentiment
Network
PW Combined PR 0.001
PW Combined PR 0.003
PW Combined PR 1.0
PW Combined PR 100.0
PW Combined PR 1000.0
Area under the curve
1.0
0.9
0.8
0.7
0.6
0.5
0.05
0.1
0.125
0.15
0.2
0.25
Evidence ratio