Movie Recommendations using Signal Diffusion

Deniz Ira, Jonathan Labhard,
Paul Griesser, Daniil Dmitriev

Outline

- Introduction
  - Motivation
  - Dataset
- Exploration
  - Graph Construction
  - Graph Properties
- Exploitation
  - Heat Diffusion
  - Results
- Conclusion

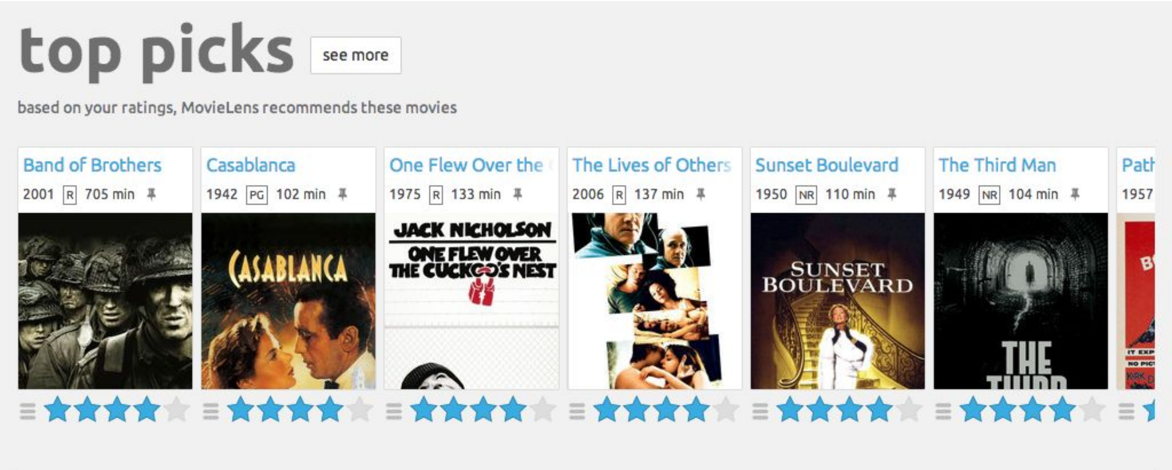
Motivations

- Recommend movies in a user-specific manner
- Explore other method than usual matrix factorization technique for system recommendation
- Use the inherent network structure represented by the data (user graph, movie graph)

Dataset

MovieLens 100k:

- 100,000 ratings 1682 movies.
- 942 users.
- Each user has rated at least 20 movies.



Constructing the graph

Based on similarities between user ratings rather than movie features.

Similarity metrics:

1. Cosine similarity "C"
2. Common ratings similarity "E"
3. Common seen movies "N"

The adjacency matrix is the weighted average of the metrics :

$$w_{i,j} = \begin{cases} 0, & \text{if } i = j \\ 1 - ((\alpha C(m_i, m_j) + \beta N(m_i, m_j) + \gamma E(m_i, m_j)/(\alpha + \beta + \gamma)) * \mathbb{I}(w_{i,j} > \epsilon) & \text{otherwise} \end{cases}$$

Hyper Parameter optimization

Grid search optimization of parameters α, β, γ and ϵ based on :

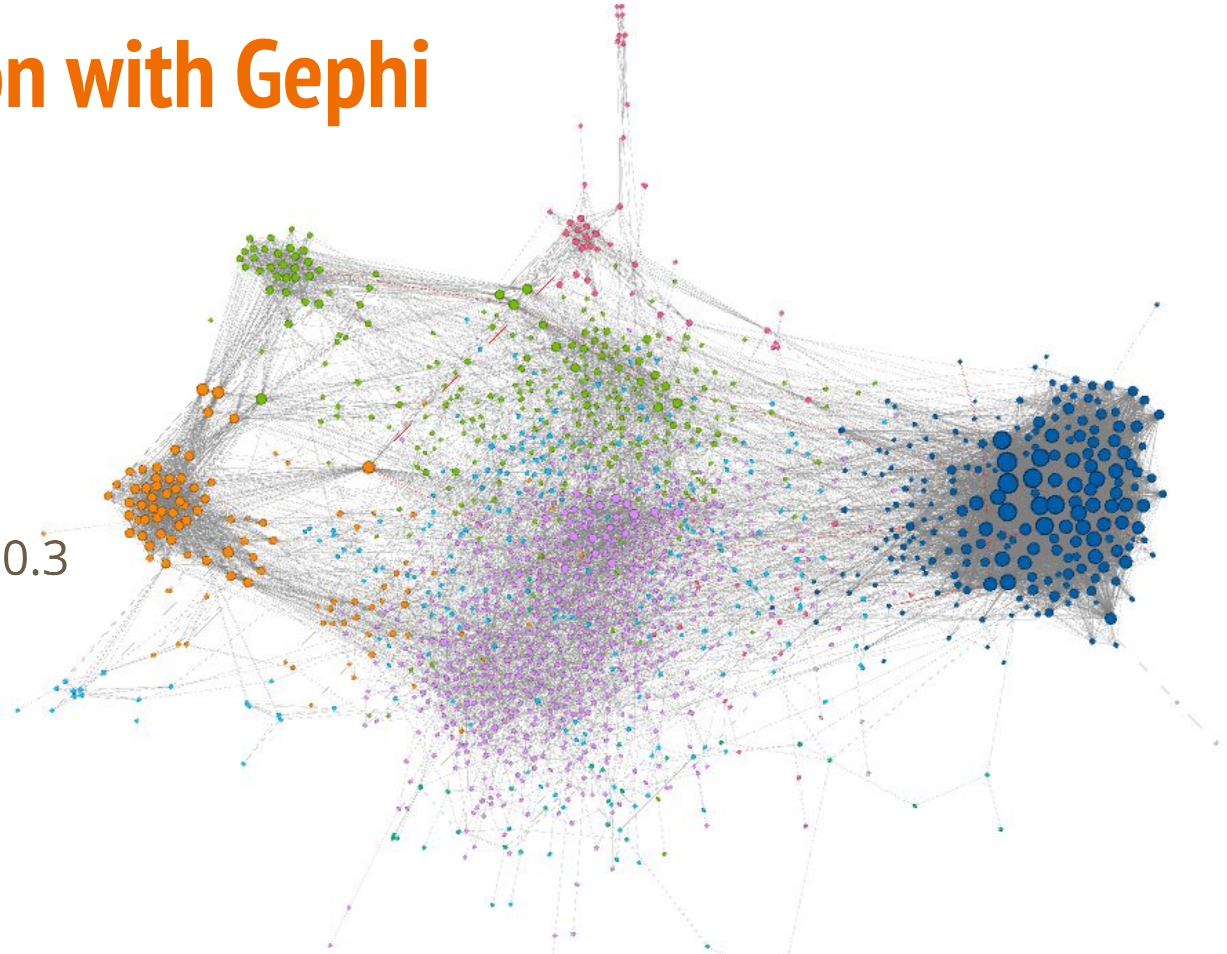
- Connected component
- Average degree
- RMSE of results

To avoid having a graph that is too connected, we remove movies with under 5 reviews.

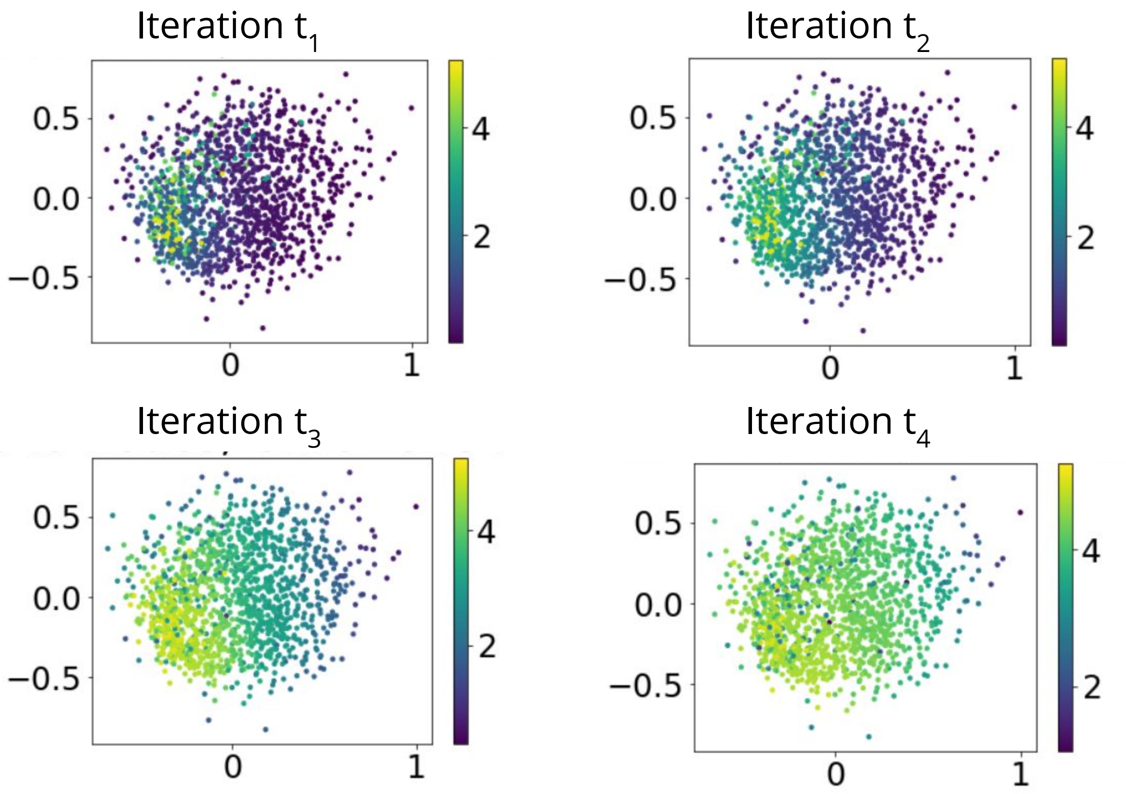
Graph visualization with Gephi

Properties:

- 1’349 nodes
- 84’841 edges
- Avg degree: 125.7
- Clustering coefficient: 0.3
- Diameter: 3



Heat Diffusion

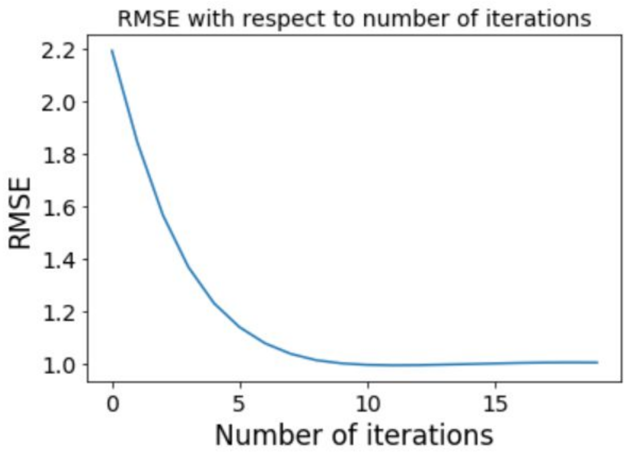


Heat Diffusion Iterations

Multiple iterations with small τ in order to have significant local information.

With each iteration, the mean rating increases.

The number of iterations is optimized to reach the best RMSE



Comparison with Baseline

Recommendations using SVD

- Schindler's List (1993)
- The Shawshank Redemption (1994)
- A Close Shave (1995)
- Psycho (1960)
- The Wrong Trousers (1993)
- The Princess Bride (1987)
- North by Northwest (1959)
- The Usual Suspects (1995)
- Wallace & Gromit (1996)
- One Flew Over the Cuckoo's Nest (1975)

Same recommendations

Hitchcock movies

Kevin Spacey movies

Recommendations using our method

- L.A. Confidential (1997)
- Pretty Woman (1990)
- Heathers (1989)
- Titanic (1997)
- Schindler's List (1993)
- Dead Man Walking (1995)
- The Shawshank Redemption (1994)
- Apocalypse Now (1979)
- Rear Window (1954)
- Get Shorty (1995)

Time for a demo!

Selected Movies

- Toy Story
- Se7en
- Apollo 13
- Star Wars
- Pulp Fiction
- Leon The Professional
- The Shawshank Redemption
- Forrest Gump
- Jurassic Park
- Blade Runner
- Silence of the Lambs
- Wallace & Gromit
- The Godfather
- 2001: A Space Odyssey
- Die Hard
- Monthy Python and the Holy Grail
- 12 Angry Men
- A Clockwork Orange
- Terminator
- The Shining
- Akira
- The Lion King

Conclusion

To go further:

- Explore other way to weight the edges in the movie graph using for example external information about the movies (genre, cast, etc)
- Explore user based graph, where nodes are user and weights between them are based on both their ratings and other informations such as hobby's or location