Alternative Similarity Functions for Graph Kernels

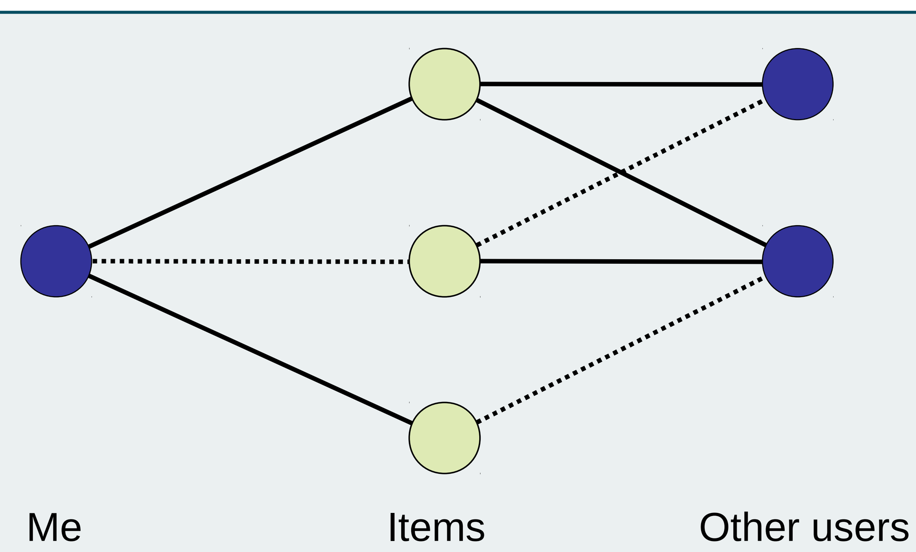
Jérôme Kunegis[1], Andreas Lommatzsch[1] & Christian Bauckhage[2]

[1]DAI-Labor, Technische Universität Berlin [2]Deutsche Telekom Laboratories

kunegis@dai-lab.de, andreas@dai-lab.de, christian.bauckhage@iais.fraunhofer.de

Abstract

Given a bipartite graph of collaborative ratings, the task of recommendation and rating prediction can be modeled with graph kernels. We interpret these graph kernels as the inverted squared Euclidean distance in a space defined by the underlying graph and show that this inverted squared Euclidean similarity function can be replaced by other similarity functions. We evaluate several such similarity functions in the context of collaborative item recommendation and rating prediction, using the exponential diffusion kernel, the von Neumann kernel, and the random forest kernel as a basis. We find that the performance of graph kernels for these tasks can be increased by using these alternative similarity functions.



Bipartite rating graph with signed edge weights

Collaborative Filtering with Graph Kernels

- Predict ratings for User×Item pair (u,i)
- Given the bipartite rating graph with symmetric adjacency matrix A
- Predict ratings by weighted mean of known ratings (user-based) using similarity function *sim(u,v)*

$$P(u,i) = (\Sigma_v \, sim(u,v))^{-1} \, \Sigma_v \, sim(u,v) \, R_{vi}$$

- Define graph kernel K_A (forest, exp., von Neuman)

$K_{FOR} = (I + D - A)^{-1}$ (forest kernel)

$K_{EXP} = \exp(\alpha A) = \Sigma_{i=0}^{\infty} \alpha^i / i! \, A^i$ (exponential kernel)

$K_{NEU} = (I - \alpha A)^{-1} = \Sigma_{i=0}^{\infty} \alpha^i A^i$ (von Neumann kernel)

- Instead of using $K_A(u,i)$ as similarity, use distance

$$d(u,v) = K_A(u,u) + K_A(v,v) - K_A(u,v) - K_A(v,u)$$

Similarity Functions

- Use similarity functions based on d(u,v)

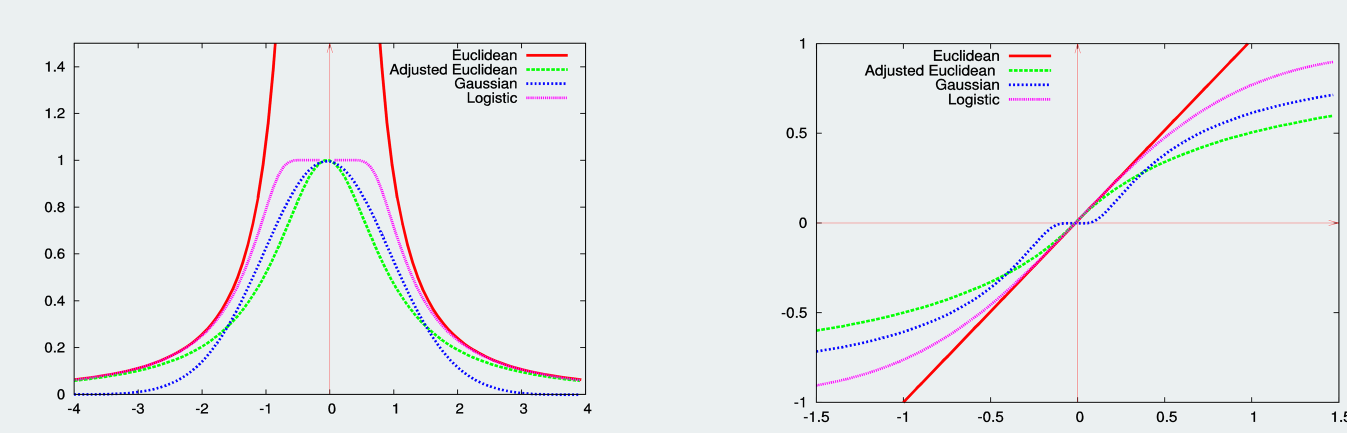
$sim_{Eu} = \sigma^2 / d^2$ *(inverted Euclidean distance)*

$sim_{EuA} = 1 / (1 + d^2 / \sigma^2)$ *(adjusted Euclidean distance)*

$sim_{Ga} = \exp(-½ \, d^2 / \sigma^2)$ *(Gaussian kernel)*

- Parameter σ^2 denotes variance
- For small σ^2, adjusted Euclidean corresponds to unadjusted Euclidean

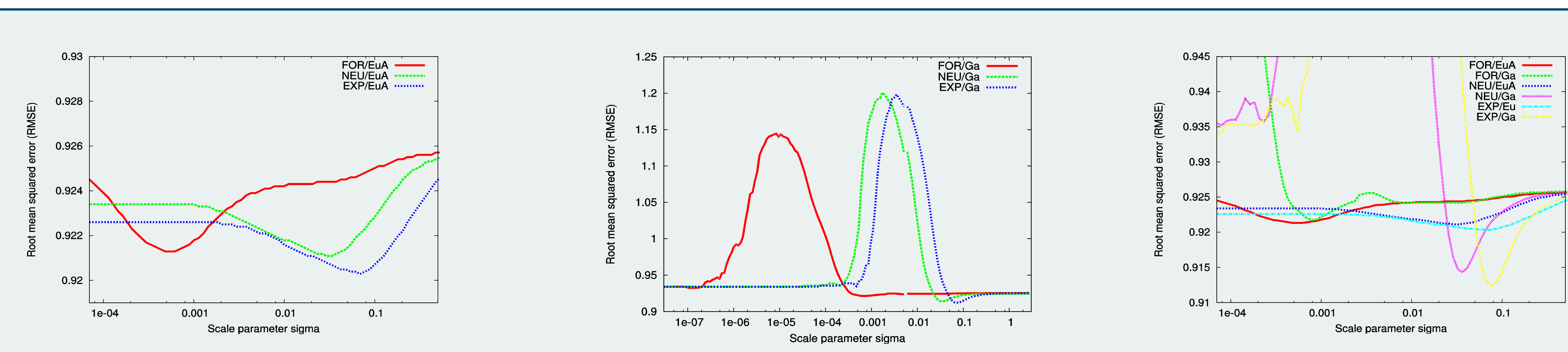
Similarity functions in function of *d* and *1/d*



(a) d (b) 1/d

Evaluation

- Predict ratings on subset of Netflix Prize Corpus
- 3,216 users, 1,307 movies and 57,507 ratings
- Compute the Root mean squared error (RMSE)



(a) Euclidean (b) Gaussian (c) Overall

References

[1] *Learning Semantic Similarity*, J. Kandola et al., KDD 2003

[2] *Diffusion Kernels on Graphs and Other Discrete Structures, R. Kondor et al., ICML 2002*

Observations

- Best prediction accuracy for specific value of σ^2
- Bad performance for adjacent values of σ^2 with Gaussian kernel
- Overall best is Gaussian similarity function