Context-aware similarities within the factorization framework

Balázs Hidasi

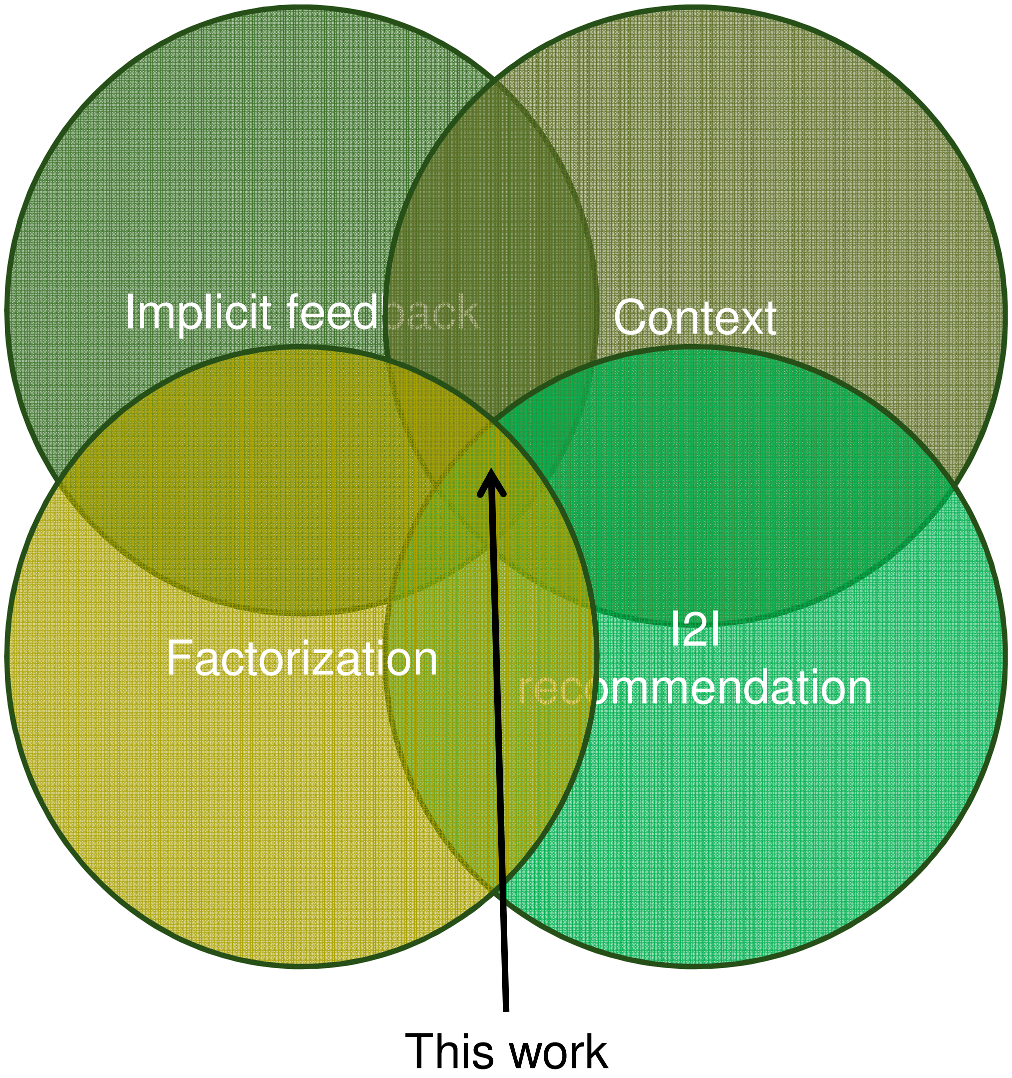
Domonkos Tikk

CaRR workshop, 5th February 2013, Rome

Outline

- Background & scope
- Levels of CA similarity
- Experiments
- Future work

BACKGROUND


Implicit feedback
Context
Factorization
I2I
recommendation
This work

IMPLICIT FEEDBACK

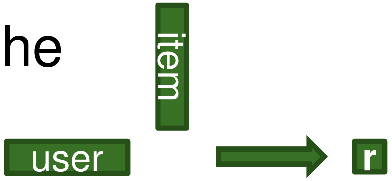
- User preference not coded explicitely in data
- E.g. purchase history
  - Presence → preference assumed
  - Absence → ???
- Binary „preference” matrix
  - Zeroes should be considered
- Optimize for
  - Weighted RMSE
  - Partial ordering (ranking)

Context

- Can mean anything
- Here: event context
  - User U has an event
  - on Item I
  - while the context is C
- E.g.: time, weather, mood of U, freshness of I, etc.
- In the experiments:
  - Seasonality (time of the day or time of the week)
    - Time period: week / day

FACTORIZATION I

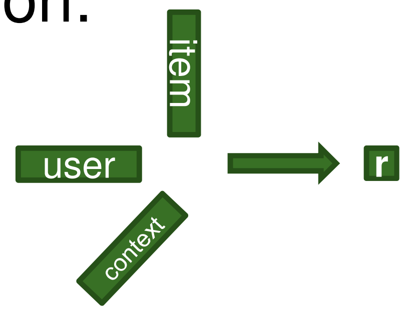
- Preference data can be organized into matrix
  - Size of dimensions high
  - Data is sparse
- Approximate this matrix by the product of two low rank matrices
  - Each item and user has a feature vector
  - Predicted preference is the scalar product of the appropriate vectors
  - $r_{u,i} = (U_u)^T I_i$

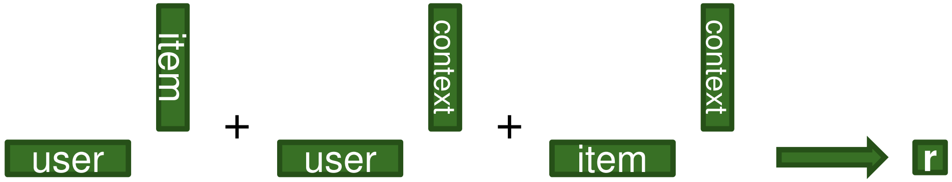


- Here we optimize for wRMSE (implicit case)
  - Learning features with ALS

FACTORIZATION II.

- Context introduced → additional context dimension
- Matrix → tensor (table of records)
- Models for preference prediction:
  - Elementwise product model
    - Weighted scalar product
    - $r_{u,i,c} = 1^T (U_u \circ I_i \circ C_c)$
  - Pairwise model
    - Context dependent user/item bias
    - $r_{u,i,c} = (U_u)^T I_i + (U_u)^T C_c + (I_i)^T C_c$





Item-to-item recommendations

- Items similar to the current item
- E.g.: user cold start, related items, etc.
- Approaches: association rules, similarity between item consumption vectors, etc.
- In the factorization framework:
  - Similarity between the feature vectors
  - Scalar product: $s_{i,j} = (I_i)^T I_j$
  - Cosine similarity: $s_{i,j} = \frac{(I_i)^T I_j}{\|I_i\|_2 \|I_j\|_2}$

SCOPE OF THIS PROJECT

- Examination wether item similarities can be improved
  - using context-aware learning or prediction
  - compared to the basic feature based solution
- Motivation:
  - If factorization models are used anyways, it would be good to use them for I2I recommendations as well
- Out of scope:
  - Comparision with other approaches (e.g. association rules)

CONTEXT-AWARE SIMILARITIES: LEVEL1

- The form of computing similarity remains
  - $s_{i,j} = (I_i)^T I_j$ $$s_{i,j} = \frac{(I_i)^T I_j}{\|I_i\|_2 \|I_j\|_2}$$
- Similarity is NOT CA
- Context-aware learning is used
  - Assumption: item models will be more accurate
  - Reasons: during the learning context is modelled separately
- Elementwise product model
  - Context related effects coded in the context feature vector
- Pairwise model
  - Context related biases removed

Context-aware similarities: level2

- Incorporating the context features
- Elementwise product model
  - Similarities reweighted by the context feature
  - Assumption: will be sensitive to the quality of the context
  - $s_{i,j,c} = 1^T \left(I_i \circ C_c \circ I_j \right)$ $\quad s_{i,j,c} = \dfrac{1^T \left(I_i \circ C_c \circ I_j \right)}{\left\| I_i \right\|_2 \left\| I_j \right\|_2}$
- Pairwise model
  - Context dependent promotions/demotions for the participating items
  - Assumption: minor improvements to the basic similarity
  - $s_{i,j,c} = \left(I_i \right)^T I_j + \left(I_i \right)^T C_c + \left(I_j \right)^T C_c$

Context-aware similarities: level2 normalization

- Normalization of context vector
- Only for cosine similarity
- Elementwise product model:
  - Makes no difference in the ordering
  - Recommendation in a given context to a given user
- Pairwise model
  - Might affect results
  - Controls the importance of item promotions/demotions

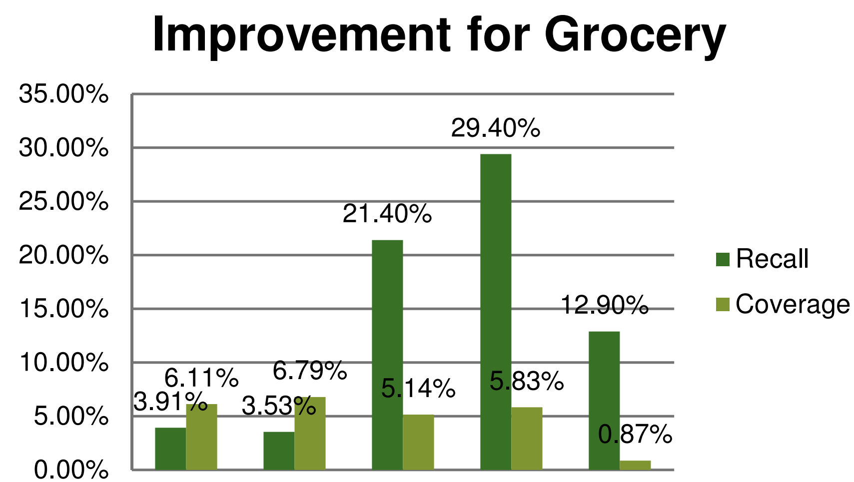
$$s_{i,j,c} = \frac{(I_i)^T I_j}{\|I_i\|_2 \|I_j\|_2} + \frac{(I_i)^T C_c}{\|I_i\|_2} + \frac{(I_j)^T C_c}{\|I_j\|_2}$$

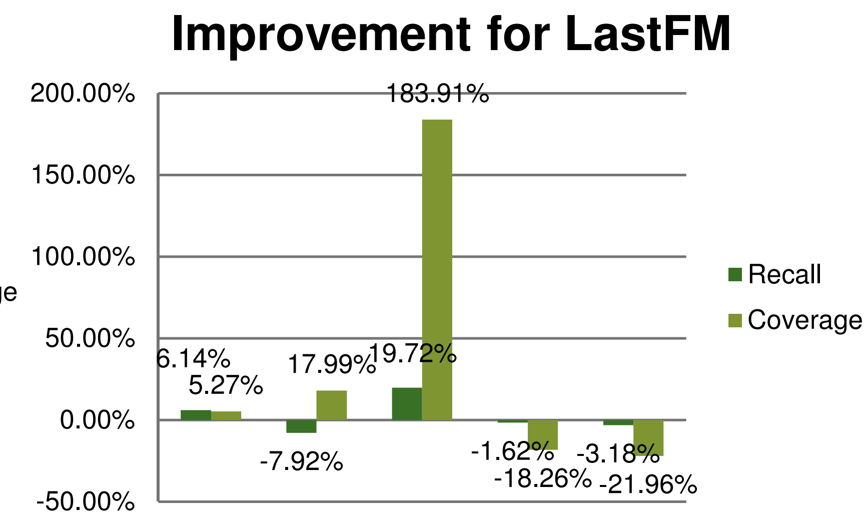
$$s_{i,j,c} = \frac{(I_i)^T I_j}{\|I_i\|_2 \|I_j\|_2} + \frac{(I_i)^T C_c}{\|I_i\|_2 \|C_c\|_2} + \frac{(I_j)^T C_c}{\|I_j\|_2 \|C_c\|_2}$$

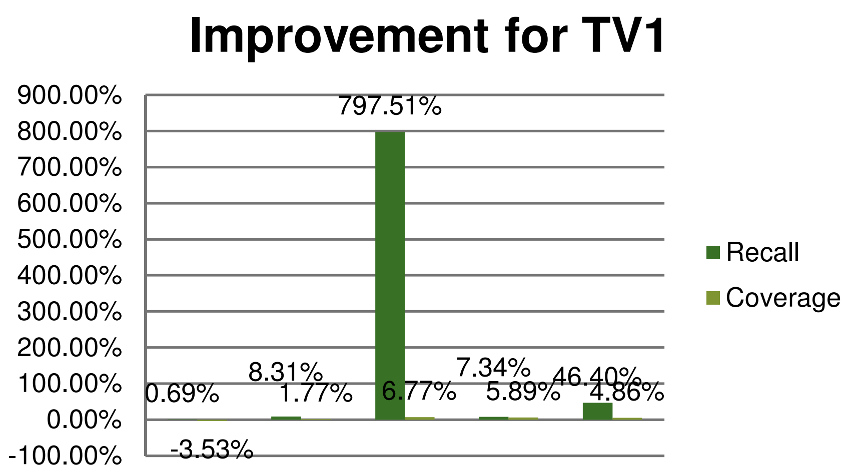
EXPERIMENTS - SETUP

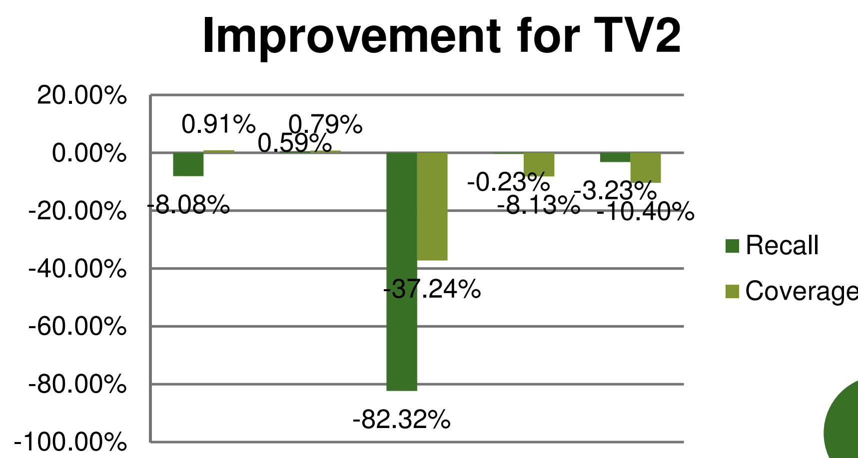
- Four implicit dataset
  - LastFM 1K – music
  - TV1, TV2 – IPTV
  - Grocery – online grocery shopping
- Context: seasonality
  - Both with manually and automatically determined time bands
- Evaluation: recommend similar items to the users' previous item
  - Recall@20
  - MAP@20
  - Coverage@20

EXPERIMENTS - RESULTS









- From left to right: L1 elementwise, L1 pairwise, L2 elementwise, L2 pairwise, L2 pairwise (norm)

EXPERIMENTS - CONCLUSION

- Context awareness generally helps
- Impromevent greatly depends on method and context quality
- All but the elementwise level2 method:
  - Minor improvements
  - Tolerant of context quality
- Elementwise product level2:
  - Possibly huge improvements
  - Or huge decrease in recommendations
  - Depends on the context/problem

(Possible) future work

- Experimentation with different contexts
- Different similarity between feature vectors
- Predetermination wether context is useful for
  - User bias
  - Item bias
  - Reweighting
- Predetermination of context quality
- Different evaluation methods
  - E.g. recommend to session

Thanks for the attention!

For more of my recommender systems related research visit my website:
http://www.hidasi.eu