Context-aware Preference Modeling with Factorization

Balázs Hidasi

RecSys'15, Doctoral Symposium

19. September 2015

GRAVITY
Research & Development

Outline

- I. Background
  - a. Implicit feedback
  - b. Context
  - c. Factorization
- II. Finished research
  - a. Context-aware algorithms (iTALS, iTALSx)
  - b. Speeding-up ALS
  - c. General Factorization Framework
- III. Future research
  - a. Automatic preference model learning
  - b. Context-related research

GRAVITY Research & Development

Background

GRAVITY
Research & Development

Implicit feedback

- + The practical scenario
- + Collected by passive monitoring
- + Available in large quantities
- - Preferences are not explicit
- - Noisy positive feedback
- - No negative feedback
- - Missing feedback needs to be handled

GRAVITY Research & Development

Context

- Context: Additional side information that can help refining the recommendations and tailoring them in order to fit the users' actual needs better.
- Context helps:
  - Dealing with context related effects during training
  - Adapting recommendation lists during recommendation time
- Types
  - User side information: user metadata, social networks, etc.
  - Item side information: item metadata, etc.
  - Context of transactions: time, location, device, etc.

Factorization

- Project entities into a low dimensional latent feature space
- The interaction between the representations estimate the preferences

Research

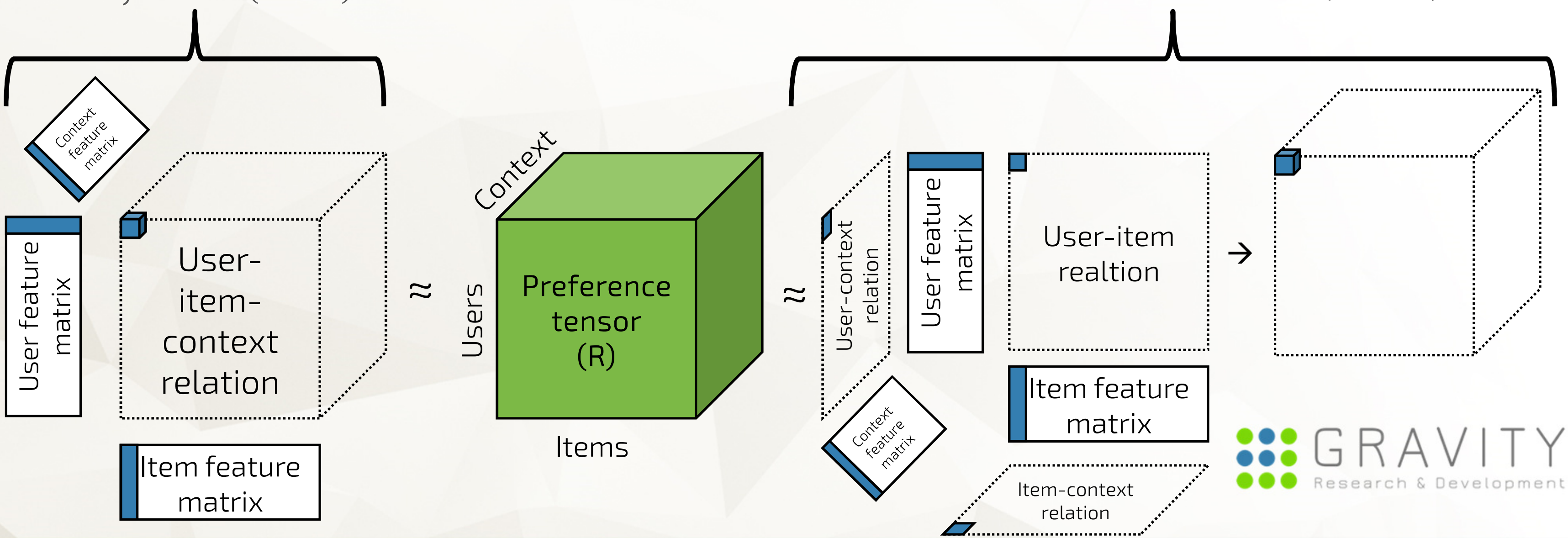
GRAVITY
Research & Development

Context-aware algorithms [1,2]

- iTALS / iTALSx
  - Pointwise preference estimation
  - ALS learning
  - Scales linearly with the number of transactions
  - Different models
- Models for different problems
  - Low number of features, sparser data → iTALSx
  - Denser data, using higher number of features is possible → iTALS

N-way model (iTALS)

Pairwise interaction model (iTALSx)



GRAVITY Research & Development

Speeding up ALS [3]

- ALS scales cubically (quadratically in practice) with the number of features
  - Bottleneck: solving a $K \times K$ system of linear equations
  - Highly impractical to use high factor models
- Approximate solutions for speed-up
  - ALS-CG: conjugate gradient based direct approximation of ALS
    - Efficiency depends on matrix-vector multiplication
  - ALS-CD: optimize on a feature-by-feature basis (instead of computing whole feature vectors)
    - Implicit case: lots of negative examples → compression

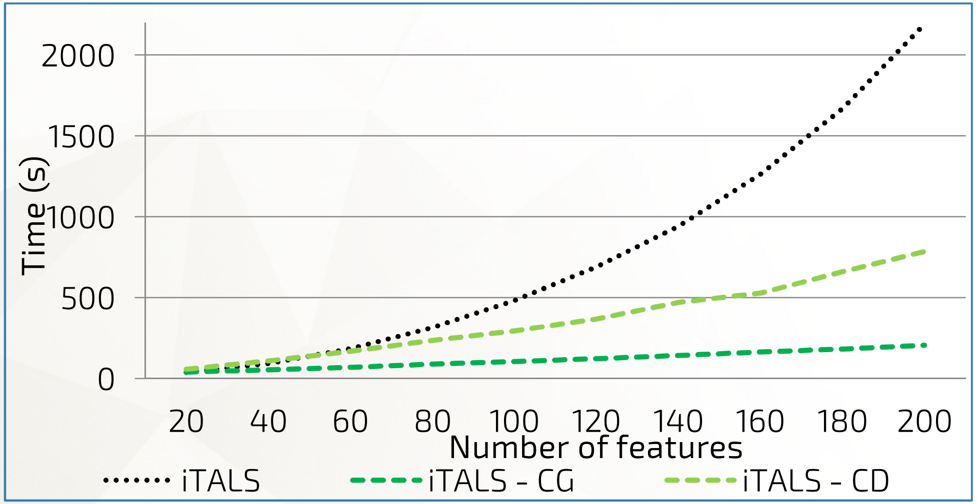
GRAVITY Research & Development

Speed-up results

- Accuracy similar to ALS

Method	Similar	Worse	Better
ALS-CG	62 of 75 (82.67%)	10 of 75 (13.33%)	3 of 75 (4%)
ALS-CD	57 of 75 (76%)	16 of 75 (21.33%)	2 of 75 (2.67%)

- Significant speed-up
  - Better trade-offs (accuracy vs. time)
  - More efficient resource usage
- Linear scaling with the number of features (in practice)
  - High factor models are usable
- CG or CD?



GFF: General Factorization Framework [4]

- *An algorithm that allows experimentation with novel models for the context-aware recommendation problem, that are not restricted to the two main model classes used by the state-of-the-art.*
- Motivation
  - N_D dimensions → lots of different possible preference models
  - Standard models not necessarily fit the problem (e.q. asymmetry)
  - Lack of tool that has this flexibility
- Features
  - No restriction on the context
  - Large preference model class
  - Data type independence
  - Flexibility
  - Scalability

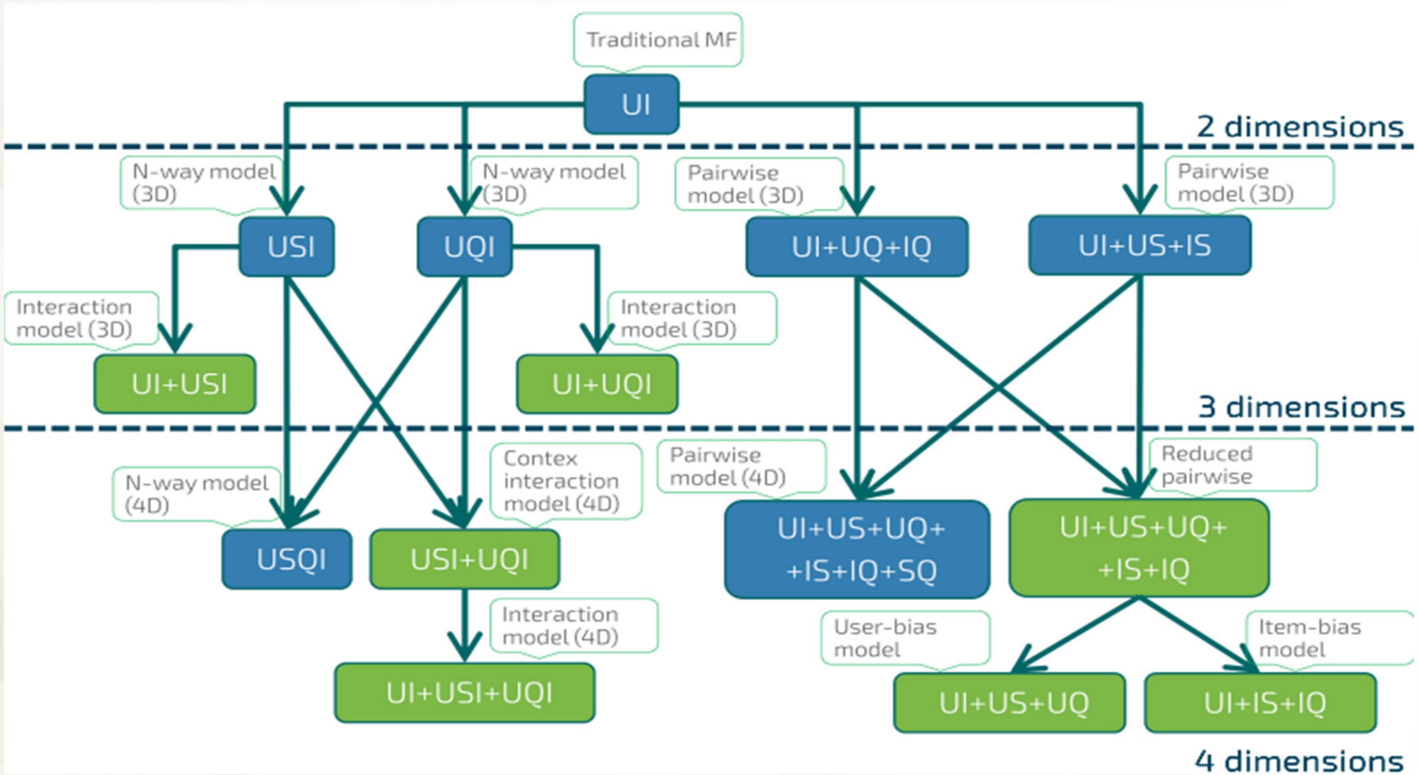
GRAVITY Research & Development

Novel preference models with GFF (1)

- Interactions with context
  - User-item
  - User-context-item (reweighting)
  - User-context (bias)
  - Item-context (bias)
  - Context-context?
- A 4D problem
  - Users (U)
  - Items (I)
  - Seasonality (S)
  - Sequentiality (Q)
- Traditional models
  - N-way (USQI)
  - Pairwise (UI+US+IS+UQ+IQ+SQ)
- Novel models
  - Interaction (UI+USI+UQI)
  - Context-interaction (USI+UQI)
  - Reduced pairwise (UI+US+IS+UQ+IQ)
  - User bias (UI+US+UQ)
  - Item bias (UI+UQ+IQ)
  - (Other interesting ones: UI+USQI; UI+USI+UQI+USQI; USI+UQI+USQI)

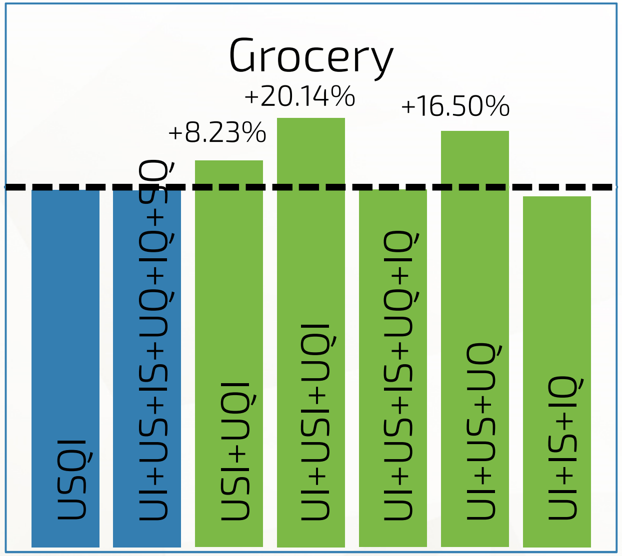
GRAVITY Research & Development

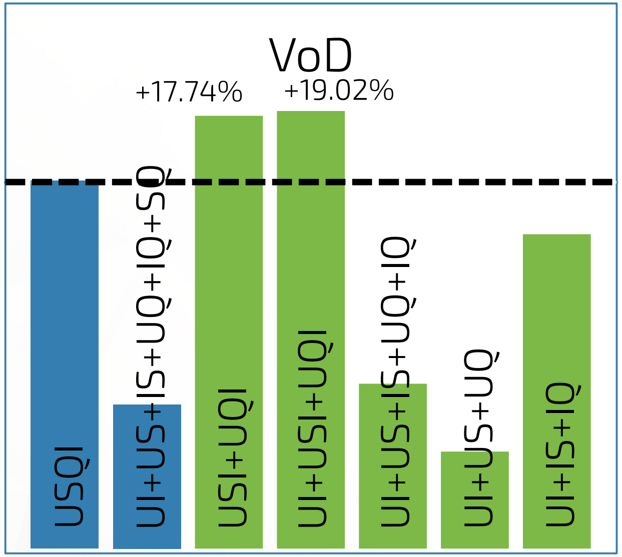
Novel preference models with GFF (2)



GRAVITY
Research & Development

Performance of novel models





Dataset	Best model	Improvement (over traditional)	Novel better than traditional
Grocery	UI+USI+UQI	+20.14%	3 of 5
TV1	USI+UQI	+15.37%	2 of 5
TV2	UI+USI+UQI	+30.30%	4 of 5
LastFM	UI+USI+UQI	+12.40%	3 of 5
VoD	UI+USI+UQI	+19.02%	2 of 5

GRAVITY Research & Development

Future research

GRAVITY
Research & Development

Automatic model learning for GFF

- Flexibility of GFF
  - Useful for experimentation
  - Finding the best (or fairly good) model requires lots of experiments for a new setup
- Automatize model selection
  - Which contexts should be used?
  - Which interactions should be used?

GRAVITY
Research & Development

Model selection with LARS

- Model: UI+US+IS+USI+UQ+IQ+UQI+USQI+USQ+ISQ+SQ
- Each term contributes to the prediction of the preferences
- Terms are the features
- Inferred preferences (0/1) are the target
  - For every possible (u,i,s,q) combination
  - Weighting: multiply examples of positive feedback by the weight

GRAVITY Research & Development

Efficiency of the model selection

- Lot of examples → efficiency?
- Efficient LARS implementations require only the
  - Covariance of features
  - Correlation of features with the target
- E.g.: $\sum_{u,i,s,q} w_{u,i,s,q} 1^T (U_u \circ S_s) 1^T \left(U_u \circ I_i \circ Q_q\right)$
  - Sum has many members
  - Can be computed efficiently
    $O\left(N^+ K^2 + S_U K^2 + S_I K^2 + S_S K^2 + S_Q K^2\right)$
  - Precomputed covariance matrices and sums of vectors required

Interaction of dimensions

- When to use the model selection?
- Dimension interact
  - One ALS epoch modifies a certain feature to be optimal with the current model
  - Different terms optimize for different aspects (e.g. USI and IS)
  - Shared features will be suboptimal to either but may lean to one side
    - Problems with unbiased selection
- Handle terms or groups of terms separately
  - Hard to integrate into solution
  - Requires multiple instances of feature matrices
  - Increases model complexity

Selection strategies

- Joint pretraining (few epochs), model selection, training selected model
- Multiple iterations of pretraining and selecting
- Joint training of a few terms, extend to full model using the trained features, (additional training), selection, train
- Separate training, model selection, (merge separate feature matrices for the same dimension), (training)
- Separate training, model selection, train non selected members on the residual

Context-related research

- Non-conventional context
  - Standard context: entity based
  - Other types
    - Hierarchical
    - Composite
    - Ordered
    - Continuous
- Context quality
  - General quality
  - Suitability for a model or interaction type
  - Improving quality by splitting/combining context-states

GRAVITY
Research & Development

Thank you!

References (papers can be downloaded from http://hidasi.eu)
[1] Balázs Hidasi and Domonkos Tikk: *Fast ALS-based tensor factorization for context-aware recommendation from implicit feedback*. ECML-PKDD (2012)
[2] Balázs Hidasi: *Factorization models for context-aware recommendations*. Infocommunications Journal VI(4) (2014)
[3] Balázs Hidasi and Domonkos Tikk: *Speeding up ALS learning via approximate methods for context-aware recommendations*. Knowledge and Information Systems (2015)
[4] Balázs Hidasi and Domonkos Tikk: *General factorization framework for context-aware recommendations*. Data Mining and Knowledge Discovery (2015)

GRAVITY
Research & Development