Context-aware Preference Modeling with Factorization

GRAVITY Research & Development
crowdrec
Balázs Hidasi
balazs.hidasi@gravityrd.com
MŰEGYETEM 1782
BME TMIT
balazs.hidasi @gravityrd.com
http://hidasi.eu

Context-aware factorization on implicit feedback [1,2]

Implicit feedback

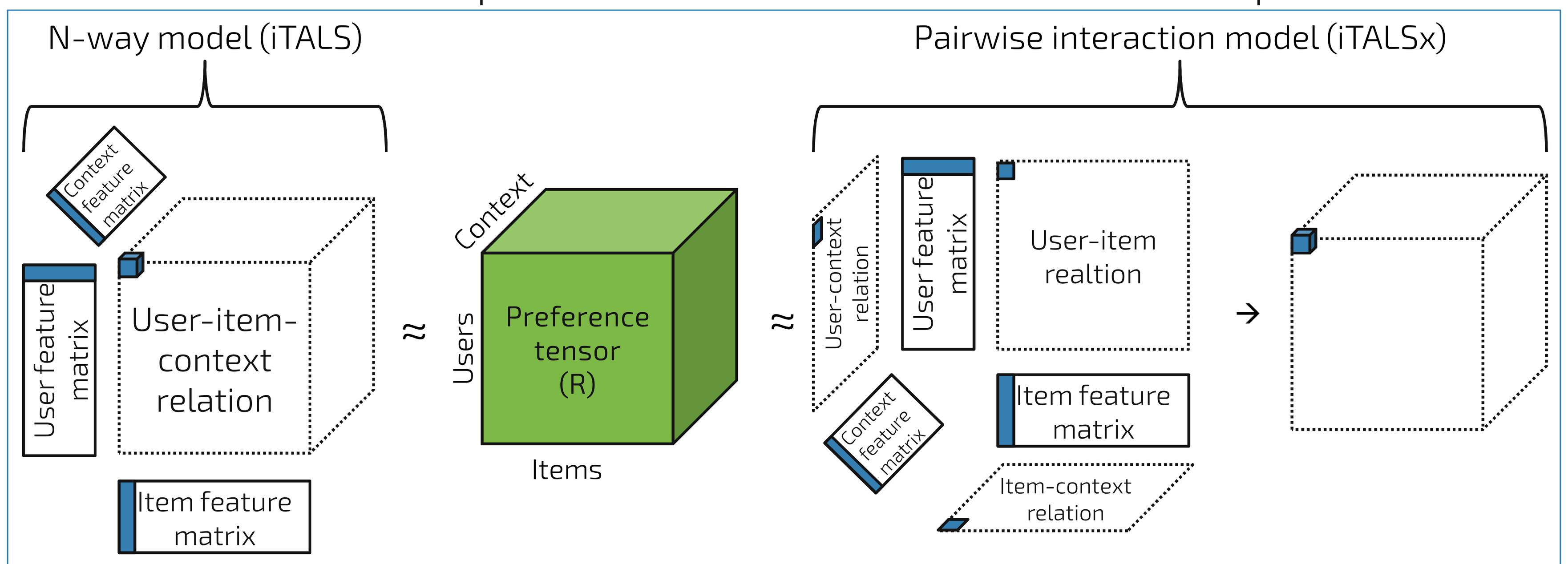
- Observation of user activity, inferred feedback
- The practical scenario (explicit feedback is scarcely available in practice)
- Noisy positive feedback
- Missing negative feedback

Context

- Any information besides the user-item transactions
- Property of the transaction rather than that of users or items
- Helps recommendations:
  - (1) Noise filtering during training (context-related patterns)
  - (2) Recommendation lists are adapted to fit the users' needs better

iTALS/iTALSx factorization algorithms

- Context-aware data organized into a tensor
- Poitwise preference estimation
  - User-item-context(s) combination in training data: 1 as preference value
  - Missing (negative) feedback: 0 as preference value
- Weight function: lower weights for missing feedback
- Loss: weighted sum of squared errors
- Optimization: Alternating Least Squares (ALS)
  - One matrix is computed at a time, the others are fixed
  - Smart decomposition of calculations allows efficient computations



Speeding up ALS learning [3]

Scaling of ALS

- Number of transactions: *linear*
- Number of features: cubical / practically *quadratic*

Approximations

- ALS-CD (Coordinate Descent – learn one feature at a time)
- ALS-CG (Conjugate Gradient – approximate the LS solution)

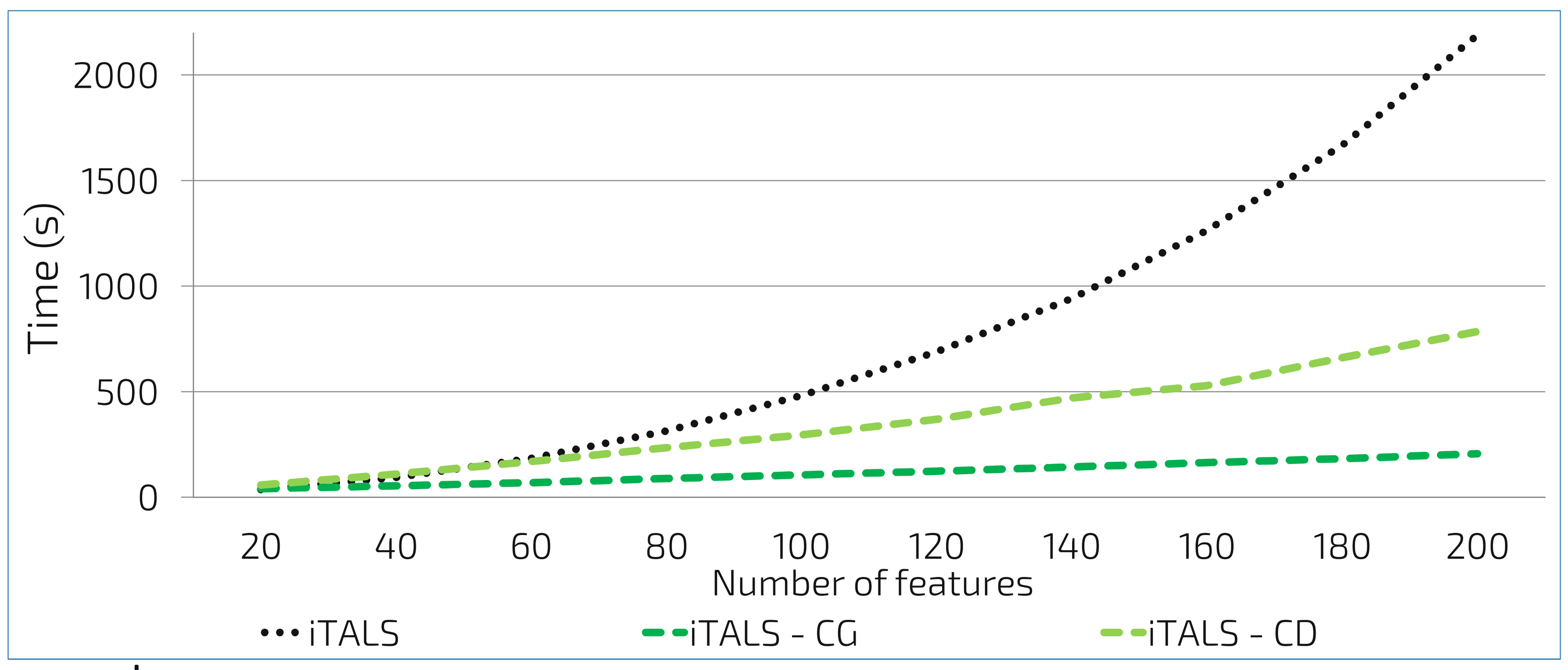
Practically *linear scaling* in the number of features (for commonly used values)

- High factor models can be used
- More frequent retrains
- Better trade-offs between accuracy and training times

Accuracy is 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%)

Considerable speed-up (up to 10 times)



Comparison

- CG is better: faster, more stable, a little bit more accurate and directly approximates the ALS solution

General Factorization Framework – experimentation with preference models [4]

Aims & Goals

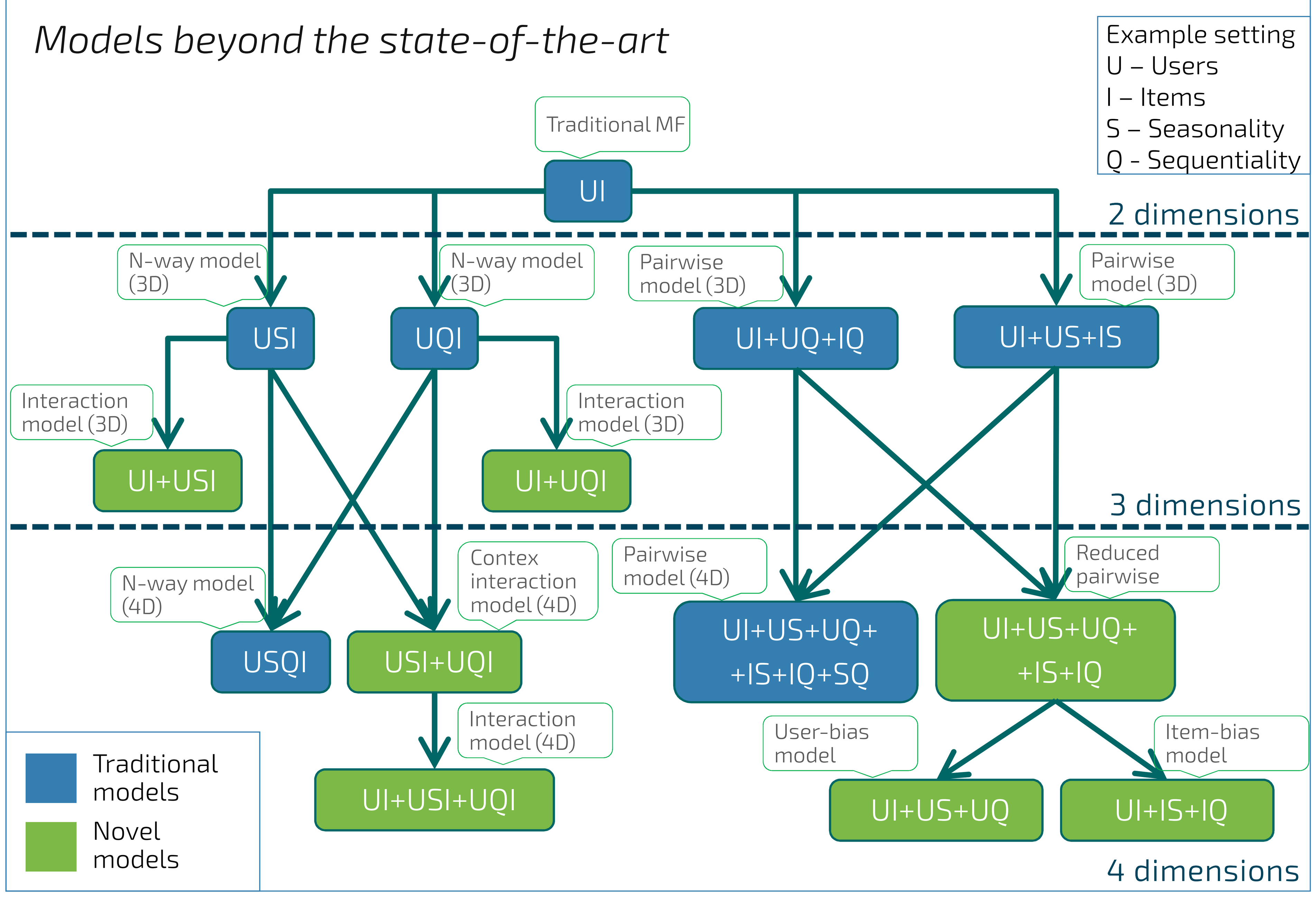
Create an algorithm that allows experimentation with novel models for the context-aware recommendation problem, that are not restricted to the two main model classes used by the state-of-the-art.

Typical interactions

User-Item: Core interaction.
User-Item-Context: Context-dependent reweighting of interactions.
User-Context: Context-dependent user bias.
Item-Context: Context-dependent item bias.
Context-Context: Required for symmetric models.

Traditional models are symmetric, recommendation problems are not.

Models beyond the state-of-the-art



Features

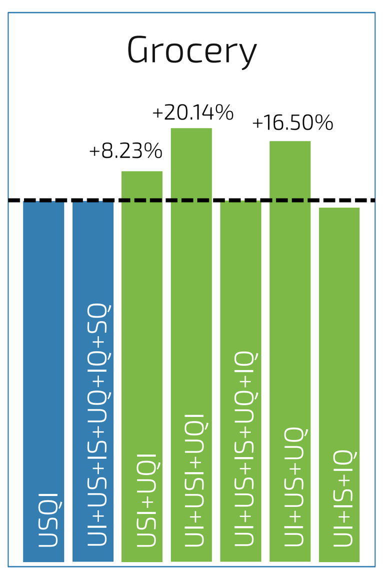
No restriction on the context
Based on the SA-MDM data model, GFF works on any context-aware recommendation problem independently of the number and the meaning of context dimensions. The extended is fully compatible with MDM, enabling the usage of additional data (e.g. session information, item metadata, etc.).

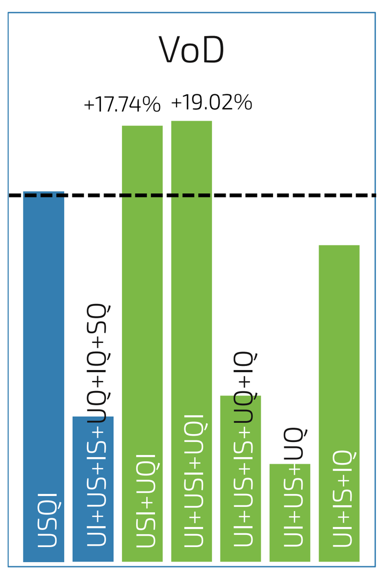
Large preference model class
The only restriction on the preference model is that it must be linear in the dimensions of the problem. (Meaning that a dimension can not directly interact with itself in the model.) This intuitive restriction does not restrict the applicability to real-world problems.

Data type independence
Besides the practically more useful implicit case, explicit problems can be also addressed by changing the weighting scheme in the loss function.

Flexibility
The weighting scheme of GFF is very flexible, enabling to incorporate extra knowledge through the weights such time decay, dwell time dependent weighting, missing not at random hypotheses and more.

Scalability
GFF scales well both in terms of the number of interactions in the training set and in the number of features. This makes it applicable in real life recommender systems.





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

[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)