Approximate modeling of continuous context in factorization algorithms

Balázs Hidasi

Domonkos Tikk

CaRR workshop, 13th April 2014, Amsterdam

CONTEXT – TO BE ON THE SAME PAGE

- Event context (transaction context)
  - Associated with the transaction:
    - Time
    - Device
    - External parameter matched to event by time (e.g. weather)
    - Etc.
  - Determined by both the item and user
    - Previously bought item
- Atomic context
  - Context value is atomic

FACTORIZATION

- Entity set (=dimension) ←→ feature matrix
  - Entity ←→ feature vector
    - Mapping to the K dimensional feature space
- Entity combination: (at most) one entity from each dimension
  - Prediction for an entity combination/interaction:
  - By an expression using their feature vectors
- Example: Matrix Factorization (MF)
  - Entity sets: users, items
  - Prediction for user's preference on item: $\hat{r}_{u,i} = U_u^T I_i$

CONTEXT WITH FACTORIZATION

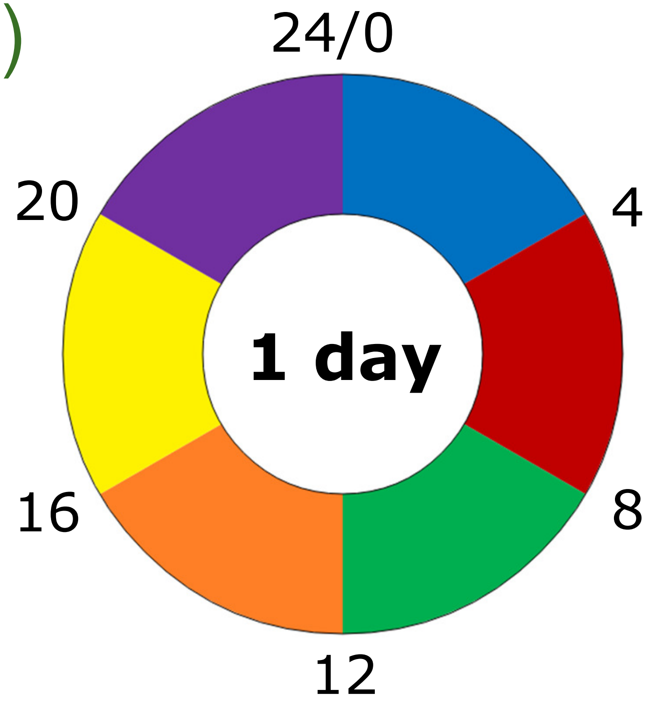
- Factorization: entity based
- Context must be:
  - Atomic
  - Nominal (categorical)
- „Entitization” of non-nominal context
  - 1. Discretization (if continuous)
  - 2. Dismissal of ordering

Diversity in context

- Previously bought item (sequentiality): nominal
- Weather (e.g.: sunny, rainy, etc): nominal?
- Screen size (of device): ordinal
  - Ordering is dismissed
- Temperature: continuous
  - Discretization: from freezing to extremely hot
  - Ordering is dismissed
- Time: continuous

Seasonality (example context dimension)

- Periodic, time based, widely used context
- Length of season (e.g. 1 day)
  - Time MOD length of season
    - Same context if time between events is N*(length of season)
- „Entitized” seasonality
  - 1. Discretization: creation of time band within the season
    - E.g.: night (0-4), dawn (4-8), morning (8-12), afternoon (12-16), late afternoon (16-20), evening (20-24)
  - 2. Dismissal of ordering:
    - Sequence and neighboring properties of time bands does not matter
    - Time bands fully independent from each other



Problems of „entitized" continuous context

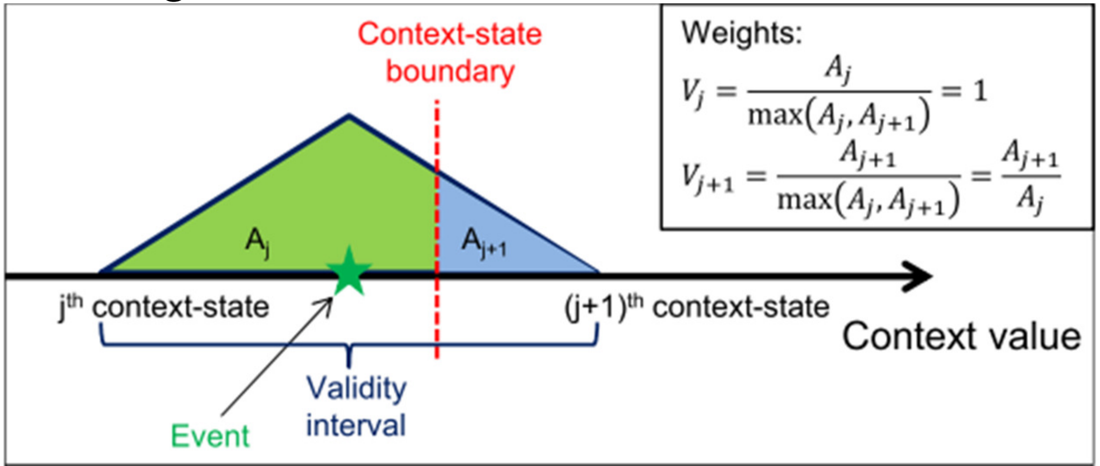
- 1. Context-state rigidness
  - Strict boundaries
  - Events close to the boundary → fully associated with one context-state
  - In reality: events close to boundary belong to both context-states to some extent
- 2. Context-state ordinality
  - Context-states treated independently
  - In reality: behavior in neighboring context-states is more similar
    - Gradual change in behavior
    - Order of context-states therefore matter
- 3. Context-state continuity
  - The representative of the context-state should be continuous
  - Stricter version of (2)

APPROXIMATE MODELING OF CONTINUOUS CONTEXT

- Modeling approaches
  - How to handle these dimensions?
  - Not a concrete algorithm
  - Can be integrated into most factorization algorithm
    - (One example with iTALS is shown)
- Why is approximate in the title?
  - Not fully continuous modeling
  - (3) is not addressed
  - Some form of discretization is kept

Approach 1: Fuzzy event modeling (1/2)

- Addresses the rigidness problem (1)
- Initial discretization
- Events near context-state boundaries
  - Associated with both
  - Event validity: instantaneous → interval
    - Event associated with every context that intersects with its interval
  - Weights for the extent association

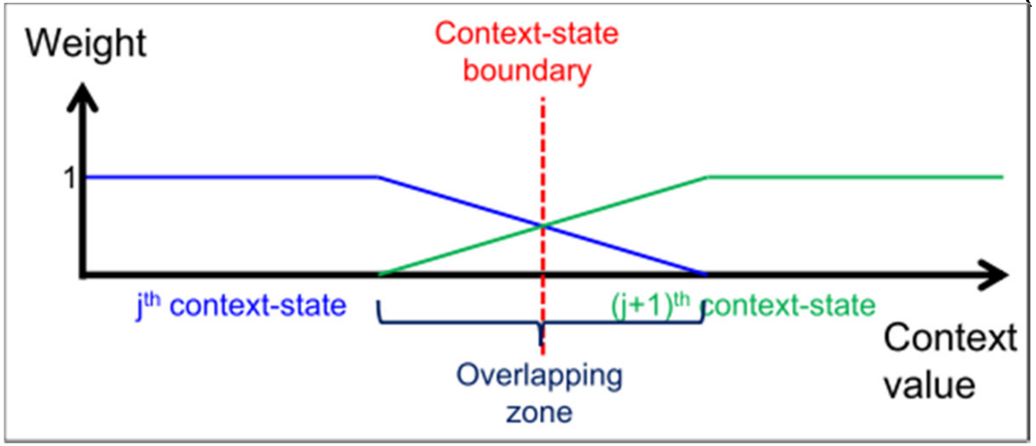


Approach 1: Fuzzy event modeling (2/2)

- Integration to factorization algorithm:
  - Duplication of certain events
  - No modification in algorithm
- Ordinality problem addressed indirectly:
  - Duplicate events are used to train neighboring context-states → enforce similarity (to some extent)

APPROACH 2: FUZZY CONTEXT MODELING

- Addresses rigidness & ordinality problem (1 & 2)
- Overlapping of context-state intervals
  - Initial discretization
  - Overlapping around the boundary → solves (1)
- Context-state of an event
  - Mixture of context-states (generally) → solves (2)
    - Here: mixture of 2 if in overlapping zone; or 1 if not
  - Weights of context-states in mixture
- Feature vector of an event for the context dimension
  - Linear combination of 2 context-state feature vectors (if in the overlapping zone)
  - The feature vector of the current context-state (otherwise)



APPLYING CONTEXT MODELING APPROACH

- Base context dimension
  - Discretization of the continuous context dimension with higher resolution than before
  - Needed, because factorization methods are entity based
- Prediction should be given using the base context dimension
- Changing prediction model
  - In place of the context-state feature vector use the mixture
- Loss should sum over the base context dimension
- For periodic context: last and first context-state are neighbors

iTALS

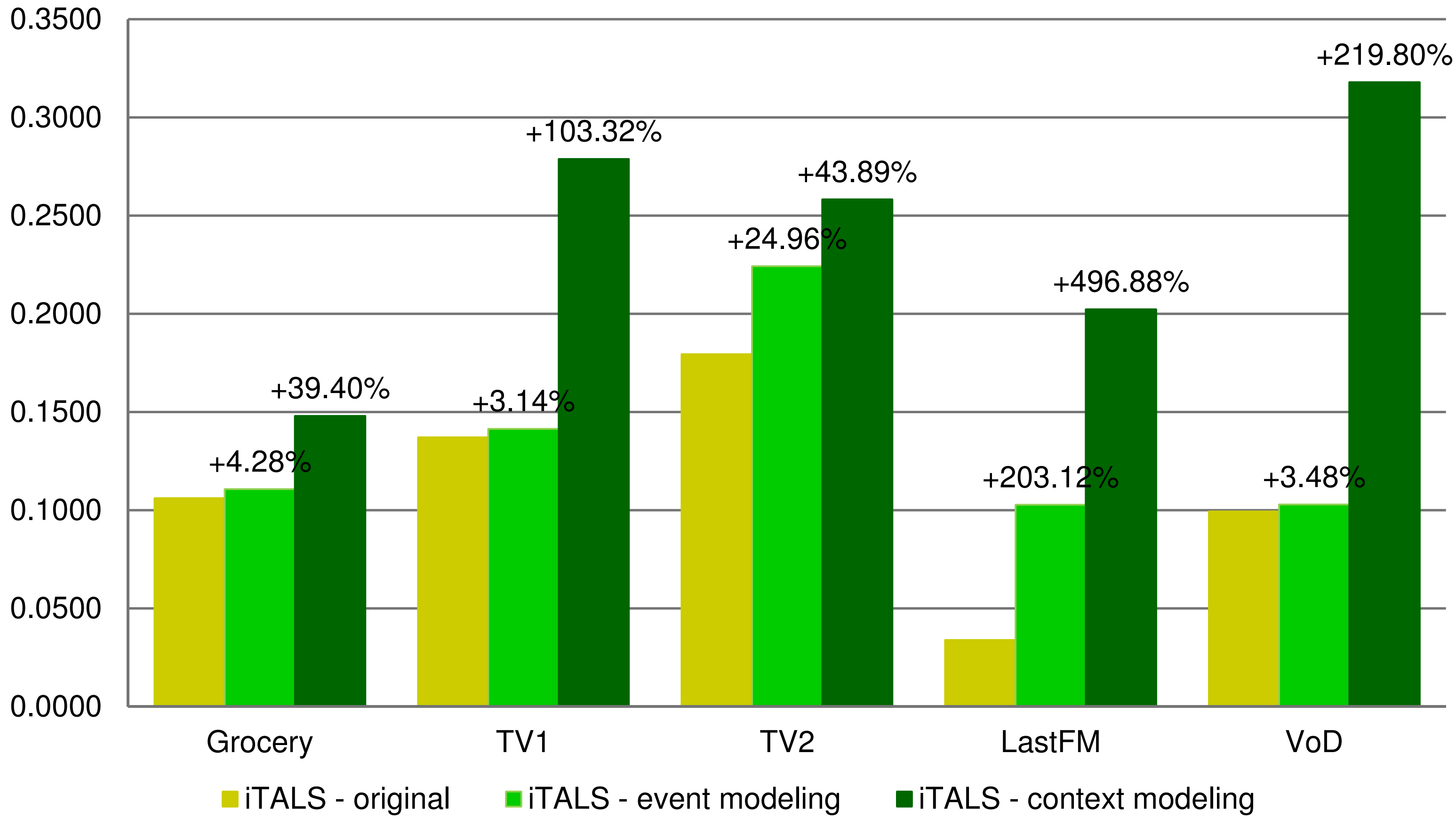
- Context-aware factorization method
  - Developed for the implicit feedback problem
  - Can be used for explicit tasks as well
- Entity sets: users, items, context(s)
  - 3+ dimensional tensor
- Prediction model: N-way model
  - $\hat{r}_{u,i,c} = 1^T (U_u \circ C_c \circ I_i)$
- Loss: Weighted Root Mean Squared Error (wRMSE)
  - $L = \sum_{u=1,i=1,c=1}^{S_U,S_I,S_C} w_{u,i,c} \left(\hat{r}_{u,i,c} - r_{u,i,c}\right)^2$
  - $w_{u,i,c} = 1$, if (u, i, c) is not in the training data
  - $w_{u,i,c} \gg 1$, if (u, i, c) is in the training data
- Learning strategy: Alternating Least Squares (ALS)

iTALS with fuzzy context modeling

- Prediction model:
  - $\hat{r}_{u,i,x} = 1^T\left(U_u \circ \left(\alpha(x)C_c + \left(1-\alpha(x)\right)C_{c+1}\right) \circ I_i\right)$
  - x: entity in the base context
  - c: entity in the original discretization of the context
  - x is a subinterval of c
- Loss function:
  - $L = \sum_{u=1,i=1,c=1}^{S_U,S_I,S_C} \sum_{x\in c} w_{u,i,x}\left(\hat{r}_{u,i,x} - r_{u,i,x}\right)^2$
- Effects on learning:
  - Computation of feature vector for c depends on $(c-1)$ and $(c+1)$ → ordinality
  - No separate computation of features
  - (Cyclical) block tridiagonal system
    - Blocks: symmetric positive definite, feature number sized matrices
    - Can be solved efficiently

Results

- 5 implicit datasets
- Recommendation accuracy, measured by recall@20
- Context: seasonality (4 hours of a day)



Summary & future work

- Problems with „entitization" of continuous context in factorization algorithms
  - (1) Rigidness problem
  - (2) Ordinality problem
- Fuzzy event modeling (for (1), indirectly for (2))
  - Events with validity interval
  - Duplication of events
  - Straightforward integration into algorithms
- Fuzzy context modeling (for (1) & (2))
  - Context-states overlap
  - Mixture context-state feature vectors
  - Considerable improvement in recommendation accuracy
- Future work
  - Applying modeling approaches to other algorithms
  - Solving the continuity problem

THANKS FOR THE ATTENTION!

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