COT: Contextual Operating Tensor for Context-aware Recommender Systems

Qiang Liu  Shu Wu  Liang Wang

Center for Research on Intelligent Perception And Computing (CRIPAC)
National Lab of Pattern Recognition (NLPR)
Institute of Automation, Chinese Academy of Sciences (CASIA)

@AAAI 2015, Austin, Jan 28, 2015
Information Overload
Context-awareness

School: An Inconvenient Truth
Home: The Lord Of The Rings

Weekdays: Data Mining
Weekends: Gone With the Wind

Child: Finding Nemo
Girlfriend: Titanic

Happy: The Merchant of Venice
Sad: Hamlet

(Time)
(Location)
Related Works

• Multiverse recommendation$^1$:  
  \[ \text{user} \times \text{item} \times \text{context} \]

• Factorization machine (FM)$^2$:  
  \[ \text{user} \times \text{item} + \text{user} \times \text{context} + \text{item} \times \text{context} \]

• Contexts are treated as other dimensions similar to the dimensions of users and items.

• Calculate the similarity among user, item and context.

  \[ \text{user} \times \text{context} \quad \times \]

  \[ \text{item} \times \text{context} \quad \times \]

(1) Karatzoglou et al, Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering.
(2) Rendle et al, Fast context-aware recommendations with factorization machines.
Ideas

(user, context) \times (item, context)
Overview of COT

Contextual operating tensors and latent vectors of entities are shown on the left side, and the computational procedure under each context combination is illustrated in the square.
Overview of COT

Matrix Factorization with Biases:

\[ \hat{r}_{i,j,k} = \omega_0 + \omega_i + \omega_j + \sum_{m=1}^{n} \omega_{m,k} + u_{i,k}^T v_{j,k} \]
Overview of COT

Contextual Operating Matrix:

\[ u_{i,k} = M_{U,k} u_i \]
\[ v_{j,k} = M_{V,k} v_j \]

Contextual operating tensor for users

Contextual operating tensor for items

Contextual operating matrix for users

Contextual operating matrix for items

Context-specific latent vector of the user

Context-specific latent vector of the item

Predicted rating

Latent vector of a user

Latent vector of an item

Weighted combination

Context operation
Overview of COT

Combination of Contexts:

\[ a_k = H_k W \]
Overview of COT

Contextual Operating Tensor:

\[ M_{U,k} = a_k^T T_{U}^{[1:d]} \]
\[ M_{V,k} = a_k^T T_{V}^{[1:d]} \]

[Diagram showing the relationship between contextual operating tensors and latent vectors]
Overview of COT

Overall Function:

\[
\hat{r}_{i,j,k} = \omega_0 + \omega_i + \omega_j + \sum_{m=1}^{n} \omega_{m,k} + \left( H_k W \right)^T T_U \left[ u_i \right] \left( H_k W \right)^T T_V \left[ v_j \right]
\]
Experiments

<table>
<thead>
<tr>
<th>compared methods</th>
<th>metrics</th>
<th>dataset splitting</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVD++</td>
<td>RMSE</td>
<td>All Users</td>
</tr>
<tr>
<td>Multiverse recommendation FM</td>
<td>MAE</td>
<td>Cold Start</td>
</tr>
<tr>
<td>HeteroMF(^1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>dataset</th>
<th>#contexts</th>
<th>contexts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food dataset</td>
<td>2</td>
<td>virtuality, hunger</td>
</tr>
<tr>
<td>Adom dataset</td>
<td>5</td>
<td>when, where, companion, release, recommendation</td>
</tr>
<tr>
<td>Movielens-1M</td>
<td>2</td>
<td>hour in a day, day in a week</td>
</tr>
</tbody>
</table>

(1) Jamali and Lakshmanan, Heteromf: recommendation in heterogeneous information networks using context dependent factor models.
## Performance Comparison

<table>
<thead>
<tr>
<th>Dataset</th>
<th>All Users</th>
<th>Cold Start</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>MAE</td>
</tr>
<tr>
<td><strong>Food Dataset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVD++</td>
<td>1.155</td>
<td>0.948</td>
</tr>
<tr>
<td>Multiverse</td>
<td>1.063</td>
<td>0.841</td>
</tr>
<tr>
<td>FM</td>
<td>1.055</td>
<td>0.845</td>
</tr>
<tr>
<td>HeteroMF</td>
<td>1.072</td>
<td>0.862</td>
</tr>
<tr>
<td>COT</td>
<td><strong>1.002</strong></td>
<td><strong>0.792</strong></td>
</tr>
<tr>
<td><strong>Adom Dataset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVD++</td>
<td>2.782</td>
<td>2.093</td>
</tr>
<tr>
<td>Multiverse</td>
<td>1.833</td>
<td>1.383</td>
</tr>
<tr>
<td>FM</td>
<td>1.852</td>
<td>1.446</td>
</tr>
<tr>
<td>HeteroMF</td>
<td>2.084</td>
<td>1.552</td>
</tr>
<tr>
<td>COT</td>
<td><strong>1.726</strong></td>
<td><strong>1.367</strong></td>
</tr>
<tr>
<td><strong>Movielens-1M</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVD++</td>
<td>0.942</td>
<td>0.721</td>
</tr>
<tr>
<td>Multiverse</td>
<td>0.883</td>
<td>0.669</td>
</tr>
<tr>
<td>FM</td>
<td>0.878</td>
<td>0.672</td>
</tr>
<tr>
<td>HeteroMF</td>
<td>0.902</td>
<td>0.686</td>
</tr>
<tr>
<td>COT</td>
<td><strong>0.841</strong></td>
<td><strong>0.645</strong></td>
</tr>
</tbody>
</table>

### Graphs

- **Food Dataset**
- **Adom Dataset**
- **Movielens-1M**
Weights of Different Contexts

Food Dataset

Adom Dataset

Movielens-1M
We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.
Distributed Representation of Contexts

We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.
Distributed Representation of Contexts

We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.
Distributed Representation of Contexts

We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.
Distributed Representation of Contexts

We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.
Distributed Representation of Contexts

We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.
Conclusion

- Model the contextual information as the semantic operation on entities.

- Use contextual operating tensor to capture the common semantic effects of contexts, and latent vectors to capture the specific properties of contexts.

- Generate the contextual operating matrix from contextual operating tensor and latent vectors.
Thanks!

Q & A

Qiang Liu
Contact: qiang.liu@nlpr.ia.ac.cn