Learning both Vertical and Horizontal Dimensions of Feature Hierarchy for Effective Recommendation

Zhu Sun, Jie Yang, Jie Zhang, Alessandro Bozzon
AAAI 2017
Main contributions

- Model two types of semantically rich feature relationships in horizontal dimension, i.e., *alternative* and *complementary*.

- Propose a unified framework (HieVH) that *seamlessly* fuses *vertical* and *horizontal* dimensions of feature hierarchy for effective recommendation.
Intuition

- A user prefers athletic style: tend to buy more Athletic clothing, e.g., athletic shoes and pants to match each other, instead of Fashion clothing, e.g., heels or skirts.

- Relationships of sibling/cousin features
  - Athletic and Fashion: alternative
  - Shoes and pants: complementary
  - Shoes and heels: alternative
Definition

• Item Relationships

Item $i$ and $j$ are alternative iff $P(e_i | e_j) < P(e_i)$ and $P(e_j | e_i) < P(e_i)$.

Item $i$ and $j$ are complementary iff $P(e_i | e_j) > P(e_i)$ and $P(e_j | e_i) > P(e_i)$. 
Measure feature relationships

- Item Co-occurrence
  
  \[ IC(i, j) = \frac{P(e_i \cap e_j)}{P(e_i) \times P(e_j)} \]

  [Theorem]
  
  i and j are alternative iff \( IC(i, j) < 1 \)
  
  i and j are complementary iff \( IC(i, j) > 1 \)

- Feature relationships

  \[ FR(f, g) = \frac{1}{|f| \times |g|} \sum_{i \in f} \sum_{j \in g} IC(i, j) \]

  \( FR(f, g) < 1 \): alternative

  \( FR(f, g) > 1 \): complementary
Empirical study

Distribution of feature relationships ($\log(FR)$) in 3 layers (Amazon Clothing dataset).

Features have alternative relationships at Layer 3 (men vs. women); Feature relationships are evenly distributed at Layer 2; most features are complementary at Layer 1.
HieVH Framework

• The basic recommendation model

\[ J = \sum_{o_{ui} \neq 0} C(r_{ui}, \langle \theta_u, \bar{\theta}_i \rangle) + \alpha \sum_{f, g \in F} \Psi(\theta_f, \theta_g) + \Omega(\Theta) \]
HieVH Framework

- Modeling Vertical Dimension: the latent factor of item is adapted by adding to it the latent factors of its affiliated features in the hierarchy

\[
\bar{\theta}_i = \Phi(\theta_i, \theta_f, \vartheta_f) = \theta_i + [\vartheta_{f1}, \vartheta_{f2}, \ldots, \vartheta_{fL}] \left[\begin{array}{c}
-\theta_{f1} \\
-\theta_{f2} \\
\vdots \\
-\theta_{fL}
\end{array}\right]_{L \times d}
\]
HieVH Framework

- Modeling Horizontal Dimension: if two features are alternative, the distance of their latent factors should be large; if complementary, the distance should be small.

\[
\Psi(\theta_f, \theta_g) = \sum_{l=1}^{L} \sum_{f, g \in \mathcal{F}^l, f < g} \sigma_{fg} \| \theta_f - \theta_g \|_F^2
\]

Where \( \sigma_{fg} = \log(\text{FR}(f, g)) \)
Validation

- **Datasets:** Amazon Web store (Clothing, Electronics, CDs and Home)

- **Comparison Methods:** MF[1], CMF[2], FM[3], TaxMF[4], Sherlock[5], ReMF[6]

Validation

Performance of comparative methods measured by AUC.

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Cases</th>
<th>MF</th>
<th>CMF</th>
<th>FM</th>
<th>TaxMF</th>
<th>Sherlock</th>
<th>ReMF</th>
<th>HieV</th>
<th>HieVC</th>
<th>HieVH</th>
<th>Improve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clothing.</td>
<td>All Users</td>
<td>0.5455</td>
<td>0.5646</td>
<td>0.6826</td>
<td>0.6509</td>
<td>0.6747</td>
<td>0.7015*</td>
<td>0.7160</td>
<td>0.7291</td>
<td>0.7375</td>
<td>5.13%</td>
</tr>
<tr>
<td></td>
<td>Cold Start</td>
<td>0.5426</td>
<td>0.5667</td>
<td>0.6629</td>
<td>0.6493</td>
<td>0.6702</td>
<td>0.7032*</td>
<td>0.7124</td>
<td>0.7284</td>
<td>0.7352</td>
<td>4.55%</td>
</tr>
<tr>
<td>Electronic.</td>
<td>All Users</td>
<td>0.5555</td>
<td>0.5762</td>
<td>0.6839</td>
<td>0.6569</td>
<td>0.6915</td>
<td>0.7337*</td>
<td>0.7512</td>
<td>0.7672</td>
<td>0.7748</td>
<td>5.60%</td>
</tr>
<tr>
<td></td>
<td>Cold Start</td>
<td>0.5526</td>
<td>0.5735</td>
<td>0.6831</td>
<td>0.6475</td>
<td>0.6982</td>
<td>0.7305*</td>
<td>0.7474</td>
<td>0.7658</td>
<td>0.7741</td>
<td>5.97%</td>
</tr>
<tr>
<td>C. &amp; V.</td>
<td>All Users</td>
<td>0.5478</td>
<td>0.5622</td>
<td>0.6356</td>
<td>0.6905</td>
<td>0.7082</td>
<td>0.7249*</td>
<td>0.7328</td>
<td>0.7516</td>
<td>0.7660</td>
<td>4.84%</td>
</tr>
<tr>
<td></td>
<td>Cold Start</td>
<td>0.5433</td>
<td>0.5609</td>
<td>0.6231</td>
<td>0.6881</td>
<td>0.7076</td>
<td>0.7243*</td>
<td>0.7315</td>
<td>0.7514</td>
<td>0.7588</td>
<td>4.76%</td>
</tr>
<tr>
<td>H. &amp; K.</td>
<td>All Users</td>
<td>0.5420</td>
<td>0.5545</td>
<td>0.6938</td>
<td>0.6469</td>
<td>0.6938</td>
<td>0.7279*</td>
<td>0.7456</td>
<td>0.7574</td>
<td>0.7667</td>
<td>5.33%</td>
</tr>
<tr>
<td></td>
<td>Cold Start</td>
<td>0.5395</td>
<td>0.5562</td>
<td>0.6915</td>
<td>0.6511</td>
<td>0.6973</td>
<td>0.7275*</td>
<td>0.7412</td>
<td>0.7554</td>
<td>0.7650</td>
<td>5.15%</td>
</tr>
</tbody>
</table>

The incorporation of horizontal dimension enhances recommendation performance.
Validation

Percentage of complementarity $C_p$ and alternativity $A_p$ between rated and recommended items

<table>
<thead>
<tr>
<th>Method</th>
<th>Layer 1</th>
<th>Layer 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$C_p$</td>
<td>$A_p$</td>
</tr>
<tr>
<td>ReMF</td>
<td>87.62%</td>
<td>12.38%</td>
</tr>
<tr>
<td>ReV</td>
<td>88.89%</td>
<td>11.11%</td>
</tr>
<tr>
<td>ReVC</td>
<td>92.62%</td>
<td>7.38%</td>
</tr>
<tr>
<td>ReVH</td>
<td>95.65%</td>
<td>4.35%</td>
</tr>
</tbody>
</table>

By incorporating the two types of feature relationships, HieVH recommends to users more items complementing to rated items.
Thanks!
CitRec2017


- Benefit the society as a whole: many challenges
  - Spatio-temporal context
  - Cross-domain
  - Conflicting interests
  - ...