Intention-aware Sequential Recommendation with Structured Intent Transition

Haoyang Li
Tsinghua University
Recommender Systems (RS) are Ubiquitous

A day in our life with Recommender Systems

Products related to this item
Problem of Today’s RS: User Intentions

- Can RS understand the **user intentions** behind the **behaviors**?
- Behaviors are highly driven by user intentions in the real world.
Identifying user intentions exactly from one item is difficult, but may be possible from a sequence of items (behaviors).
Problem of Today’s RS: User Intentions

- Can RS model the **transition** of user intentions explicitly?
  - In the real world, user intentions could be **dynamic** rather than static.
  - “dynamic” : “intent transition”

![Diagram showing health transitioning to sports]
Sequence Modeling

- **Sequential Recommendation**
  - Order matters in real-world situations.

- **Sequence Modeling**
  - **Markov Chain**: fails on long sequences, data sparsity problem
  - **Recurrent Neural Network**: fails on longer sequences, high cost
  - **Transformer**
Modeling User Intentions is Challenging

- The existing methods capture behavior patterns. They fail to identify user intentions and model intent transition explicitly.

- The properties of user intentions in Recommendation:
  - Unobserved auxiliary information
  - Multiple identify
  - Correlated cognitive relations
  - Dynamic intent transition
Problem

- Intention-aware Sequential Recommendation
  - Goal: for each user, given the sequence of interacted items $v_1, ..., v_t$ with available description information, predict the $v_{t+1}$ at time index $t + 1$.
    - Description information: title, categories, reviews...
  - We extract keywords from the description information and refer to these extracted keywords as concepts.
    - User intentions could be reflected in these concepts.
    - E.g.,
      - “Yes, it’s perfect for self-defense.” → “perfect” “self-defense”
      - “This might be great for my son playing baseball.” → “great” “son” “baseball”
Model Framework

1. Transformer-based Encoder

2. Intent Extraction

3. Structured Intent Transition

4. Intent Decoder
Transformer-based Encoder

- Embedding Submodule
  - Embedding \( h_i \) = item \( v_i \) + position \( p_i \) + concepts \( c_j \)

  - Item Property
  - Order
  - Potential Intents

- Input hidden representation:
  \[
  H^0 = [h^0_1, h^0_2, \ldots, h^0_T] \quad h_i = v_i + p_i + \sum_{c_{i,j}=1} c_j
  \]

- Self-attention Submodule
  - Capture the dependencies among items within a behavior sequence

  \[
  S^l = SA(H^l) = \text{Attention}(H^lW^l_Q, H^lW^l_K, H^lW^l_V) \\
  H^{l+1} = \text{FFN}(S^l) = \text{ReLU}(S^lW^l_1 + b^1_1)W^l_2 + b^1_2
  \]

- Goals: Learn users’ behavior patterns and filter some noises
Intent Extraction

- In this step, we explicitly extract explainable user intents from the encoded sequence hidden representations $X = H^L$.

- Goal: Infer a multi-hot intention vector $m_t = [m_{t,1}, ..., m_{t,K}]$
  - $m_{t,k} = 1 \iff k^{th}$ concept belongs to the user intentions at time $t$

- A straightforward method: treating $m_t$ as a parameter to be optimized? (×)
  - over-parameterization; cause efficiency burdens

- We adopt a prototype-based method:
  - define $K$ intention prototypes, calculate similarity $s_{t,k} = \frac{x_t \cdot c_k}{\|x_t\|_2 \|c_k\|_2}$
  - draw $m_t$ from $m_t \sim \text{Categorical}(\text{Softmax}(s_{t,1}, s_{t,2}, ..., s_{t,K}))$
Structured Intent Transition

- In this step, we model intent transitions with GNN.
  - Learn a personalized intent feature matrix $Z_t = [z_{t,1}, \ldots, z_{t,K}] \in \mathbb{R}^{K \times d'}$
  - From item space to intention space $z_{t,k} = m_{t,k} \text{MLP}_k(x_t)$
  - Model the intent transition on the concept graph $A$ (ConceptNet)
  - Pre-defined concepts and their relations can be treated as knowledge; underlying cognitive activity is stored in the connections among concepts.
  - We adopt the message-passing framework $Z_{t+1} = \mathcal{F}(Z_t, A)$

Item space

Structured intent space
Intent Decoder

- After obtaining the future intent features \( Z_{t+1} \) and intent vector \( m_{t+1} \), the intent decoder is defined as:

\[
x_{t+1} = \sum_{k=1}^{K} m_{t+1,k} \text{MLP}_k'(z_{t+1,k})
\]

- Recommendation probability of item \( v_{t+1} \):

\[
p(v_{t+1}|v_1, v_2, \ldots, v_t) = \text{Softmax}(x_{t+1}V^T)
\]

- Objective Function

\[
\mathcal{L}_u = \frac{1}{|S(u)|} \sum_{v_{t+1} \in S(u)} -\log p(v_{t+1}|v_1, v_2, \ldots, v_t)
\]
Experiments

- We aim to answer the following three questions:
  - **Q1**: How does our method perform compared with other state-of-the-art sequential recommendation methods?
  - **Q2**: Can our method identify explainable user intents and model the structured intent transition accurately?
  - **Q3**: Is the intent extraction and structured intent transition module helpful in our method?
Experimental Settings

- **Datasets**
  - **Amazon**: contains a large number of product reviews from Amazon.com
    - We choose the “Beauty” category dataset.
  - **Steam**: a popular online video game platform
  - **Epinions**: a popular online consumer review website Epinions.com
  - **MovieLens**: a dataset about movie rating, including ML-1m an ML-20m
Experimental Settings

- Datasets preprocessing procedures
  - convert all reviews/ratings to implicit feedback of 1
  - remove users and items if they have fewer than 5 records
  - build the interaction sequence sorted by the timestamps for each user
  - obtain concepts of items from the available meta-data
    - e.g., items’ descriptions, reviews, etc.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>#Users</th>
<th>#Items</th>
<th>#Interactions</th>
<th>Avg.length</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beauty</td>
<td>40,226</td>
<td>54,542</td>
<td>0.35m</td>
<td>8.8</td>
<td>0.02%</td>
</tr>
<tr>
<td>Steam</td>
<td>281,428</td>
<td>13,044</td>
<td>3.5m</td>
<td>12.4</td>
<td>0.10%</td>
</tr>
<tr>
<td>Epinions</td>
<td>5,015</td>
<td>8,335</td>
<td>26.9k</td>
<td>5.37</td>
<td>0.06%</td>
</tr>
<tr>
<td>ML-1m</td>
<td>6,040</td>
<td>3,416</td>
<td>1.0m</td>
<td>163.5</td>
<td>4.79%</td>
</tr>
<tr>
<td>ML-20m</td>
<td>138,493</td>
<td>26,744</td>
<td>20m</td>
<td>144.4</td>
<td>0.54%</td>
</tr>
</tbody>
</table>

TABLE 3: Statistics of the datasets.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>#Concepts</th>
<th>#Edges</th>
<th>Avg.concepts/item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beauty</td>
<td>592</td>
<td>2,791</td>
<td>4.45</td>
</tr>
<tr>
<td>Steam</td>
<td>229</td>
<td>472</td>
<td>4.49</td>
</tr>
<tr>
<td>Epinions</td>
<td>114</td>
<td>467</td>
<td>5.50</td>
</tr>
<tr>
<td>ML-1m</td>
<td>96</td>
<td>327</td>
<td>1.94</td>
</tr>
<tr>
<td>ML-20m</td>
<td>316</td>
<td>842</td>
<td>4.21</td>
</tr>
</tbody>
</table>

TABLE 4: Statistics of preprocessed concepts of the datasets.
Experimental Settings

- **Evaluation metrics**
  - **Hit Rate (HR)**
    \[
    \text{HR}_{@k} = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \delta(|\mathcal{T}_u \cap \mathcal{R}_{u,k}| > 0)
    \]
  - **Normalized Discounted Cumulative Gain (NDCG)**
    \[
    \text{NDCG}_{@k} = \frac{1}{Z} \text{DCG}_{@k}
    \]
    \[
    = \frac{1}{Z} \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \sum_{i=1}^{k} \frac{\delta(r_{u,i} \in \mathcal{T}_u)}{\log_2(i + 1)}
    \]
  - **Mean Reciprocal Rank (MRR)**
    \[
    \text{MRR} = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{1}{\text{rank}_{u}}
    \]
Baselines

- Non-sequential methods
  - PopRec, BPR-MF, NCF

- Sequential methods
  - Markov chain based methods: FPMC, Caser
  - RNN based methods: GRU4Rec, GRU4Rec+
  - Transformer based methods: SASRec, BERT4Rec
Experimental Results

- **Recommendation Accuracy**

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Metric</th>
<th>PopRec</th>
<th>BPR-MF</th>
<th>NCF</th>
<th>FPMC</th>
<th>GRU4Rec</th>
<th>GRU4Rec$^+$</th>
<th>DCGF</th>
<th>Caser</th>
<th>SASRec</th>
<th>BERTRec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beauty</td>
<td>HR@10</td>
<td>0.0077</td>
<td>0.0415</td>
<td>0.0407</td>
<td>0.0345</td>
<td>0.0402</td>
<td>0.0551</td>
<td>0.0626</td>
<td>0.0475</td>
<td>0.0906</td>
<td>0.0953</td>
</tr>
<tr>
<td>HR@5</td>
<td>0.0392</td>
<td>0.1209</td>
<td>0.1305</td>
<td>0.1387</td>
<td>0.1315</td>
<td></td>
<td>0.1781</td>
<td>0.1835</td>
<td>0.1625</td>
<td>0.1934</td>
<td>0.2207</td>
</tr>
<tr>
<td>HR@10</td>
<td>0.0762</td>
<td>0.1992</td>
<td>0.2142</td>
<td>0.2401</td>
<td>0.2343</td>
<td></td>
<td>0.2654</td>
<td>0.2778</td>
<td>0.2590</td>
<td>0.2633</td>
<td>0.3025</td>
</tr>
<tr>
<td>NDCG@5</td>
<td>0.0230</td>
<td>0.0814</td>
<td>0.0855</td>
<td>0.0902</td>
<td>0.0922</td>
<td></td>
<td>0.1172</td>
<td>0.1241</td>
<td>0.1050</td>
<td>0.1346</td>
<td>0.1590</td>
</tr>
<tr>
<td>NDCG@10</td>
<td>0.0349</td>
<td>0.1064</td>
<td>0.1124</td>
<td>0.1211</td>
<td>0.1074</td>
<td></td>
<td>0.1453</td>
<td>0.1543</td>
<td>0.1360</td>
<td>0.1633</td>
<td>0.1862</td>
</tr>
<tr>
<td>MR</td>
<td>0.0437</td>
<td>0.1006</td>
<td>0.1043</td>
<td>0.1056</td>
<td>0.1023</td>
<td></td>
<td>0.1299</td>
<td>0.1381</td>
<td>0.1205</td>
<td>0.1536</td>
<td>0.1701</td>
</tr>
<tr>
<td>Steam</td>
<td>HR@10</td>
<td>0.0159</td>
<td>0.0314</td>
<td>0.0246</td>
<td>0.0358</td>
<td>0.0574</td>
<td>0.0812</td>
<td>0.0564</td>
<td>0.0495</td>
<td>0.0885</td>
<td>0.0957</td>
</tr>
<tr>
<td>HR@5</td>
<td>0.0805</td>
<td>0.1177</td>
<td>0.1203</td>
<td>0.1517</td>
<td>0.2171</td>
<td></td>
<td>0.2391</td>
<td>0.1825</td>
<td>0.1766</td>
<td>0.2559</td>
<td>0.2710</td>
</tr>
<tr>
<td>HR@10</td>
<td>0.1389</td>
<td>0.1993</td>
<td>0.2169</td>
<td>0.2551</td>
<td>0.3133</td>
<td></td>
<td>0.3954</td>
<td>0.2954</td>
<td>0.2870</td>
<td>0.3783</td>
<td>0.4013</td>
</tr>
<tr>
<td>NDCG@5</td>
<td>0.0472</td>
<td>0.0744</td>
<td>0.0717</td>
<td>0.0945</td>
<td>0.1370</td>
<td></td>
<td>0.1613</td>
<td>0.1392</td>
<td>0.1131</td>
<td>0.1727</td>
<td>0.1843</td>
</tr>
<tr>
<td>NDCG@10</td>
<td>0.0665</td>
<td>0.1005</td>
<td>0.1026</td>
<td>0.1283</td>
<td>0.1802</td>
<td></td>
<td>0.2053</td>
<td>0.1717</td>
<td>0.1484</td>
<td>0.2147</td>
<td>0.2261</td>
</tr>
<tr>
<td>MR</td>
<td>0.0669</td>
<td>0.0942</td>
<td>0.0932</td>
<td>0.1139</td>
<td>0.1420</td>
<td></td>
<td>0.1757</td>
<td>0.1400</td>
<td>0.1305</td>
<td>0.1874</td>
<td>0.1949</td>
</tr>
<tr>
<td>Epinions</td>
<td>HR@10</td>
<td>0.0075</td>
<td>0.0151</td>
<td>0.0155</td>
<td>0.0162</td>
<td>0.0169</td>
<td>0.0176</td>
<td>0.0188</td>
<td>0.0164</td>
<td>0.0217</td>
<td>0.0220</td>
</tr>
<tr>
<td>HR@5</td>
<td>0.0339</td>
<td>0.0472</td>
<td>0.0538</td>
<td>0.0578</td>
<td>0.0629</td>
<td></td>
<td>0.0737</td>
<td>0.0736</td>
<td>0.0733</td>
<td>0.0822</td>
<td>0.0866</td>
</tr>
<tr>
<td>HR@10</td>
<td>0.0831</td>
<td>0.1005</td>
<td>0.0975</td>
<td>0.1083</td>
<td>0.1280</td>
<td></td>
<td>0.1380</td>
<td>0.1153</td>
<td>0.1351</td>
<td>0.1358</td>
<td>0.1462</td>
</tr>
<tr>
<td>NDCG@5</td>
<td>0.0206</td>
<td>0.0316</td>
<td>0.0338</td>
<td>0.0373</td>
<td>0.0431</td>
<td></td>
<td>0.0456</td>
<td>0.0491</td>
<td>0.0444</td>
<td>0.0530</td>
<td>0.0534</td>
</tr>
<tr>
<td>NDCG@10</td>
<td>0.0358</td>
<td>0.0464</td>
<td>0.0474</td>
<td>0.0512</td>
<td>0.0565</td>
<td></td>
<td>0.0657</td>
<td>0.0656</td>
<td>0.0642</td>
<td>0.0701</td>
<td>0.0724</td>
</tr>
<tr>
<td>MR</td>
<td>0.0430</td>
<td>0.0540</td>
<td>0.0543</td>
<td>0.0546</td>
<td>0.0581</td>
<td></td>
<td>0.0700</td>
<td>0.0693</td>
<td>0.0668</td>
<td>0.0699</td>
<td>0.0705</td>
</tr>
<tr>
<td>ML-1m</td>
<td>HR@10</td>
<td>0.0141</td>
<td>0.0914</td>
<td>0.0397</td>
<td>0.1386</td>
<td>0.1383</td>
<td>0.2092</td>
<td>0.1770</td>
<td>0.2194</td>
<td>0.2351</td>
<td>0.2863</td>
</tr>
<tr>
<td>HR@5</td>
<td>0.0715</td>
<td>0.2866</td>
<td>0.1932</td>
<td>0.4297</td>
<td>0.4673</td>
<td></td>
<td>0.3103</td>
<td>0.4485</td>
<td>0.5533</td>
<td>0.5845</td>
<td>0.6076</td>
</tr>
<tr>
<td>HR@10</td>
<td>0.1358</td>
<td>0.4301</td>
<td>0.3477</td>
<td>0.5946</td>
<td>0.6207</td>
<td></td>
<td>0.6351</td>
<td>0.6032</td>
<td>0.6629</td>
<td>0.6629</td>
<td>0.6970</td>
</tr>
<tr>
<td>NDCG@5</td>
<td>0.0416</td>
<td>0.1903</td>
<td>0.1146</td>
<td>0.2885</td>
<td>0.3196</td>
<td></td>
<td>0.3705</td>
<td>0.3162</td>
<td>0.3832</td>
<td>0.3980</td>
<td>0.4454</td>
</tr>
<tr>
<td>NDCG@10</td>
<td>0.0621</td>
<td>0.2565</td>
<td>0.1640</td>
<td>0.3439</td>
<td>0.3627</td>
<td></td>
<td>0.4064</td>
<td>0.3660</td>
<td>0.4268</td>
<td>0.4368</td>
<td>0.4818</td>
</tr>
<tr>
<td>MR</td>
<td>0.0627</td>
<td>0.2009</td>
<td>0.1358</td>
<td>0.2891</td>
<td>0.3041</td>
<td></td>
<td>0.3462</td>
<td>0.3105</td>
<td>0.3648</td>
<td>0.3790</td>
<td>0.4254</td>
</tr>
<tr>
<td>ML-20m</td>
<td>HR@10</td>
<td>0.0221</td>
<td>0.0553</td>
<td>0.0231</td>
<td>0.1079</td>
<td>0.1459</td>
<td>0.2021</td>
<td>0.1760</td>
<td>0.1232</td>
<td>0.2544</td>
<td>0.3400</td>
</tr>
<tr>
<td>HR@5</td>
<td>0.0805</td>
<td>0.2128</td>
<td>0.1358</td>
<td>0.3601</td>
<td>0.4657</td>
<td></td>
<td>0.5118</td>
<td>0.4361</td>
<td>0.3804</td>
<td>0.5727</td>
<td>0.6333</td>
</tr>
<tr>
<td>HR@10</td>
<td>0.1378</td>
<td>0.3538</td>
<td>0.2922</td>
<td>0.5201</td>
<td>0.5844</td>
<td></td>
<td>0.6524</td>
<td>0.6252</td>
<td>0.5427</td>
<td>0.7166</td>
<td>0.7473</td>
</tr>
<tr>
<td>NDCG@5</td>
<td>0.0511</td>
<td>0.1332</td>
<td>0.0771</td>
<td>0.2239</td>
<td>0.3090</td>
<td></td>
<td>0.3630</td>
<td>0.3267</td>
<td>0.2538</td>
<td>0.4208</td>
<td>0.4967</td>
</tr>
<tr>
<td>NDCG@10</td>
<td>0.0695</td>
<td>0.1786</td>
<td>0.1271</td>
<td>0.2895</td>
<td>0.3837</td>
<td></td>
<td>0.4087</td>
<td>0.3809</td>
<td>0.3062</td>
<td>0.4665</td>
<td>0.5340</td>
</tr>
<tr>
<td>MR</td>
<td>0.0709</td>
<td>0.1703</td>
<td>0.1072</td>
<td>0.2273</td>
<td>0.2967</td>
<td></td>
<td>0.3476</td>
<td>0.3278</td>
<td>0.2529</td>
<td>0.4026</td>
<td>0.4785</td>
</tr>
</tbody>
</table>

- ISRec: 0.1233, 0.2938
- Improv.: 0.2734, 0.2388
- Sequential > Non-sequential
- Self-attention can provide large performance gains.
- Our method outperforms all the baselines.
- The sparser the dataset, the larger the improvement of our method.
Experimental Results

- Showcases of Intent Extraction and Structured Intent Transition

```
Candidate Intents: 
brightening, moisturizers, defense, mousse, fiber, wrinkle

Intent Transition

Candidate Intents: 
curly, scalp, coconut, sweet, raw, bars

Intent Transition

Candidate Intents: 
triple, minerals, fiber, age, radiant, purifying, toner, skin

Intent Transition

Candidate Intents: 
nourishing, face, milk, straightener
```
Ablation studies and hyperparameters sensitivities

TABLE 5: Performance comparison of ISRec and variants.

<table>
<thead>
<tr>
<th></th>
<th>Beauty</th>
<th></th>
<th>ML-1m</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR@10</td>
<td>NDCG@10</td>
<td>HR@10</td>
<td>NDCG@10</td>
</tr>
<tr>
<td>ISRec</td>
<td>0.3594</td>
<td>0.2296</td>
<td>0.7363</td>
<td>0.5189</td>
</tr>
<tr>
<td>w/o GNN</td>
<td>0.3311</td>
<td>0.2095</td>
<td>0.7222</td>
<td>0.4978</td>
</tr>
<tr>
<td>w/o GNN&amp;Intent</td>
<td>0.3092</td>
<td>0.1965</td>
<td>0.7058</td>
<td>0.4731</td>
</tr>
<tr>
<td>BERT4Rec + concept</td>
<td>0.3037</td>
<td>0.1886</td>
<td>0.6987</td>
<td>0.4824</td>
</tr>
<tr>
<td>SASRec + concept</td>
<td>0.3061</td>
<td>0.1845</td>
<td>0.6972</td>
<td>0.4643</td>
</tr>
</tbody>
</table>
Conclusions

- We study the intent-aware sequential recommendation with structured intent transition.
- Why can’t the current Sequential Recommender Systems make us satisfied enough?

知其然，而不知其所以然

User Behaviors   User Intentions

Modeling user intentions with GNN: Try to promote human-like Recommender Systems.
References

Intention-aware Sequential Recommendation with Structured Intent Transition

Abstract:
Human behaviors in recommendation systems are driven by many high-level, complex, and evolving intentions behind their decision making processes. In order to achieve better performance, it is important for recommendation systems to be aware of user intentions besides considering the historical interaction behaviors. However, user intentions are seldom fully or easily observed in practice, so that the existing works are incapable of fully tracking and modeling user intentions, not to mention using them effectively into recommendation. In this paper, we present the Intention-Aware Sequential Recommendation (ISRec) method, for capturing the underlying intentions of each user that may lead to her next consumption behavior and improving recommendation performance. Specifically, we first extract the intentions of the target user from sequential contexts, then take complex intent transition into account through the message-passing mechanism on an intention graph, and finally obtain the future intentions of this target user from inference on the intention graph. The sequential recommendation for a user will be made based on the predicted user intentions, offering more transparent and explainable intermediate results for each recommendation. Extensive experiments on various real-world datasets demonstrate the superiority of our method against several state-of-the-art baselines in sequential recommendation in terms of different metrics.

Published in: IEEE Transactions on Knowledge and Data Engineering (Early Access)
Thanks!

Haoyang Li, Tsinghua University
lihy18@mails.tsinghua.edu.cn

Model Framework

Experim ental Settings

- Datasets
  - Amazon: contains a large number of product reviews from Amazon.com
  - We choose the “Beauty” category dataset.
  - Steam: a popular online video game platform
  - Epinions: a popular online consumer review website Epinions.com
  - MovieLens: a dataset about movie rating, including ML-1m and ML-20m

Conclusions

- Why can’t the current Sequential Recommender Systems make us satisfied enough?

知其然，但不知其所以然

User Behaviors

User Intentions

Modeling user intentions with GNN:
Try to promote human-like Recommender Systems.