**Summary**

- Challenge of typical SDS: Predefined Ontology & Hidden Semantics
  1. Predefined domain ontology is required to support corresponding functionality
     - Structured knowledge resources are available (e.g., Freebase, Wikipedia, FrameNet) and may provide semantic information
  2. Hidden semantics may contain important semantics
     - Implicit information helps infer feature relations

- Approach: Feature-Enriched MF-SLU
  - Enrich semantics with the structured knowledge for improving intent prediction
  - A single matrix integrating different-level knowledge for reasoning and prediction simultaneously

- Result
  - Feature-enriched MF-SLU benefits from hidden information and rich features, and outperforms the baseline that uses a language-modeling retrieval model.

**2. Model Learning by Matrix Factorization**

- Modeling Implicit Feedback:
  \[ f^+ \rightarrow f^+ \rightarrow f^+ = \begin{bmatrix} x^+ \end{bmatrix} \rightarrow p(f^+) > p(f^-) \]

- Objective:
  \[ \sum_{f^+ \in \mathcal{F}} \sum_{f^- \in \mathcal{F}} \ln \sigma(f_{j+} - f_{j-}) \phi_{k+} - \phi_{k-} \]

  MF learns a set of well-ranked intents per utterance.

**3. Experiments**

- **Dataset:** single-turn request with intents below
- **Evaluation Metrics**
  - Mean Average Precision (MAP)
  - Precision at K (P@K)

- **MAP for Intent Modeling**

<table>
<thead>
<tr>
<th>Feature Matrix (MAP)</th>
<th>ASR Transcripts</th>
<th>Test Utterance 1: I would like to contact alex</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM 25.1</td>
<td>MF-SLU 29.2 (+16.2%)</td>
<td>26.1 30.4 (+16.4%)</td>
</tr>
<tr>
<td>Word Observation</td>
<td>Embedding-Enriched Semantics</td>
<td>Type-Embedding-Enriched Semantics</td>
</tr>
<tr>
<td>32.0</td>
<td>34.2 (+6.8%)</td>
<td>33.3 33.3 (+0.2%)</td>
</tr>
<tr>
<td>31.5</td>
<td>32.2 (+2.1%)</td>
<td>32.9 34.0 (+3.4%)</td>
</tr>
</tbody>
</table>

Enriched semantics significantly improve the performance for intent modeling.

- **P@10 for Intent Modeling**

<table>
<thead>
<tr>
<th>Feature Matrix (P@10)</th>
<th>ASR Transcripts</th>
<th>Test Utterance 1: I would like to contact alex</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM 28.6</td>
<td>MF-SLU 29.5 (+3.4%)</td>
<td>29.2 30.1 (+2.8%)</td>
</tr>
<tr>
<td>Word Observation</td>
<td>Embedding-Enriched Semantics</td>
<td>Type-Embedding-Enriched Semantics</td>
</tr>
<tr>
<td>31.2</td>
<td>32.5 (+4.3%)</td>
<td>32.0 33.0 (+3.4%)</td>
</tr>
<tr>
<td>31.3</td>
<td>30.6 (-2.3%)</td>
<td>32.5 34.7 (+6.8%)</td>
</tr>
</tbody>
</table>

Type information inferred from ASR results may not be accurate enough; noisy enriched information could be degrading performance. When there are no recognition errors, accurate type information benefits performance.

**Conclusion**

- We propose an MF approach to learn user intents based on rich feature patterns from multiple modalities, including app descriptions, automatically acquired knowledge and user utterances.
- In a smart-phone intelligent assistant setting (e.g., requesting an app), the feature-enriched MF-SLU can handle users’ open domain intents by returning relevant apps that provide desired functionality either locally available or by suggesting installation of suitable apps in an unsupervised way.
- The framework can flexibly extend to incorporate different-level features for improving a system’s ability to assist users pursuing personalized multi-app activities.
- The effectiveness of the feature-enriched MF-SLU model can be shown for different domains, indicating good generality and provides a promising direction for future work.