Overview of the NTCIR-12 IMine-2 task

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Takehiro Yamamoto</td>
<td>Kyoto University</td>
</tr>
<tr>
<td>Yiqun Liu</td>
<td>Tsinghua University</td>
</tr>
<tr>
<td>Min Zhang</td>
<td>Tsinghua University</td>
</tr>
<tr>
<td>Zhicheng Dou</td>
<td>Renmin University of China</td>
</tr>
<tr>
<td>Ke Zhou</td>
<td>Yahoo! Labs London</td>
</tr>
<tr>
<td>Ilya Markov</td>
<td>University of Amsterdam</td>
</tr>
<tr>
<td>Makoto P. Kato</td>
<td>Kyoto University</td>
</tr>
<tr>
<td>Hiroaki Ohshima</td>
<td>Kyoto University</td>
</tr>
<tr>
<td>Sumio Fujita</td>
<td>Yahoo! JAPAN Corporation</td>
</tr>
</tbody>
</table>
iPhone 6
Goal of IMine

Understanding search intent behind the query
Given an ambiguous/underspecified query, produce a single result page that satisfies different user intents.

Search Result Diversification (INTENT, INTENT-2, IMine-1)
Key Challenges

- Identify possible **search intents** behind the query
  - review, official, photo, ...

- Generate a **diversified ranked list** that satisfies different search intents
What's new in IMine-2?

IMine-2

II

Search Result Diversification + Vertical Intents
Vertical Intents

We have to identify relevant verticals as well as search intents of a query.
SUBTASKS
Two Subtasks

- Query Understanding subtask
  \( \equiv \) Subtopic Mining subtask

- Vertical Incorporating subtask
  \( \equiv \) Document ranking subtask
Given a query, the system is required to generate a diversified ranked list of subtopics with their relevant vertical intents.

<table>
<thead>
<tr>
<th>Subtopic</th>
<th>Importance</th>
<th>Vertical Intent</th>
</tr>
</thead>
<tbody>
<tr>
<td>iPhone 6 Apple</td>
<td>0.95</td>
<td>Web</td>
</tr>
<tr>
<td>iPhone 6 sales</td>
<td>0.80</td>
<td>News</td>
</tr>
<tr>
<td>iPhone 6 image</td>
<td>0.73</td>
<td>Image</td>
</tr>
<tr>
<td>iPhone 6 review</td>
<td>0.65</td>
<td>Web</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Given a query, the system is required to generate a diversified ranked list of subtopics with their relevant vertical intents.

<table>
<thead>
<tr>
<th>Subtopic</th>
<th>Importance</th>
<th>Vertical Intent</th>
</tr>
</thead>
<tbody>
<tr>
<td>iPhone 6 Apple</td>
<td>0.95</td>
<td>Web</td>
</tr>
<tr>
<td>iPhone 6 sales</td>
<td>0.80</td>
<td>News</td>
</tr>
<tr>
<td>iPhone 6 image</td>
<td>0.73</td>
<td>Image</td>
</tr>
</tbody>
</table>

**Challenges**
- **Subtopic Relevance**
- **Subtopic Diversity**
- **Vertical Relevance**
Vertica
Given a query and the document collection, the system is required to return a diversified ranked list of documents.

<table>
<thead>
<tr>
<th>Documents</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>doc-011</td>
<td>0.95</td>
</tr>
<tr>
<td>Vertical-news</td>
<td>0.80</td>
</tr>
<tr>
<td>doc-021</td>
<td>0.73</td>
</tr>
<tr>
<td>Vertical-image</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Virtual vertical document

= virtual document that represents ideal vertical result of the query

= Participants can focus on which vertical and where it should be incorporated into the ranked list
Given a query and the document collection, the system is required to return a diversified ranked list of documents.

<table>
<thead>
<tr>
<th>Documents</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>doc-011</td>
<td>0.95</td>
</tr>
<tr>
<td>Vertical-news</td>
<td>0.80</td>
</tr>
</tbody>
</table>

**Virtual vertical document**

= virtual document that represents ideal vertical

- **Document Relevance**
- **Document Diversity**
- **Vertical Relevance**
DATA
Topics

- **100 queries** for each language
  - 50 queries in IMine-1

- **Five types**
  - Ambiguous (e.g. *jaguar*)
  - Faceted (e.g. *iPhone6*)
  - Very clear (e.g. *NII address*)
  - Task-oriented (e.g. *lose weight*)
  - Vertical-oriented (e.g., *iPhone photo*)
Data

- Document Corpus
  - for Vertical Incorporating subtask
  - 500 search results of a commercial search engine for each query

- Query suggestions
  - collected by several commercial search engines

- Sogou search user behavior data

- Yahoo! JAPAN related search query data
NII と Yahoo! JAPAN が検索技術研究のために新たな提携
「Yahoo!検索」の検索クエリデータを NII のワークショップに無償提供

ヤフー株式会社（以下 Yahoo! JAPAN、代表取締役社長：宮坂 学、東京都港区）と大学共同利用機関法人 情報・システム研究機構 国立情報学研究所（以下 NII、所長：喜連川 優、東京都千代田区）は、情報学研究の一層の推進に寄与するために、Yahoo! JAPAN より NII に「Yahoo!検索」の検索クエリデータを無償提供することになりました。本データは、情報アクセス技術の評価ワークショップ「NTCIR（エンティサイル）」(*1) の参加者向けに提供され、同ワークショップに参加する研究グループは無償で活用できます。

Yahoo! JAPAN が提供するデータは、平成 21 年（2009 年）7 月から平成 25 年（2013 年）6 月の期間に「Yahoo!検索」で検索された全クエリ（ユーザーが検索時に入力した単語やフレーズ）の中から、現在進行している NTCIR の第 12 サイクル（NTCIR-12）で設定された研究課題に対する関連度の高い
EVALUATION
Ground Truth Construction

**Intent**

1. **review**
   - **review**
   - **Quora**
   - **0.60**
   - **NonRel**

2. **accessory**
   - **case**
   - **bumper**
   - **accessories**
   - **0.30**
   - **Rel**

**Vertical Importance**

<table>
<thead>
<tr>
<th>Vertical</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web</td>
<td>0.30</td>
</tr>
<tr>
<td>Image</td>
<td>0.10</td>
</tr>
<tr>
<td>News</td>
<td>0.00</td>
</tr>
<tr>
<td>Encyclopedia</td>
<td>0.00</td>
</tr>
<tr>
<td>QA</td>
<td>0.10</td>
</tr>
<tr>
<td>Shopping</td>
<td>0.50</td>
</tr>
</tbody>
</table>

**Per-intent Relevance**

- **Intent 1:**
  - **review**
  - **Quora**
  - **0.60**
  - **NonRel**

- **Intent 2:**
  - **accessory**
  - **case**
  - **bumper**
  - **accessories**
  - **0.30**
  - **Rel**
E.g. 6,119 subtopics for Chinese subtask
<table>
<thead>
<tr>
<th>Intent2</th>
<th>accessory</th>
<th>case</th>
<th>bumper</th>
<th>accessories</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Web</th>
<th>Image</th>
<th>News</th>
<th>Encyclopedia</th>
<th>QA</th>
<th>Shopping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>

... 

|       | 0.30      | 0.10  | 0.00   | 0.00         | 0.10| 0.50     |
Per-Intent Document Relevance

E.g. 6,778 documents for Chinese subtask
### Query Understanding Subtask (again)

**INPUT** query  
**OUTPUT** subtopics and their vertical intents

<table>
<thead>
<tr>
<th>Subtopic</th>
<th>Importance</th>
<th>Vertical Intent</th>
</tr>
</thead>
<tbody>
<tr>
<td>iPhone 6 Apple</td>
<td>0.95</td>
<td>Web</td>
</tr>
<tr>
<td>iPhone 6 sales</td>
<td>0.80</td>
<td>News</td>
</tr>
<tr>
<td>iPhone 6 image</td>
<td>0.73</td>
<td>Image</td>
</tr>
<tr>
<td>iPhone 6 review</td>
<td>0.65</td>
<td>Web</td>
</tr>
</tbody>
</table>
• **Subtopic Diversity**
  
  – D# proposed by [Sakai+2011]

  \[
  D\#-nDCG = \gamma I-rec + (1 - \gamma) D-nDCG
  \]

  How many intents are covered by the subtopics?
  How much relevant are the subtopics?

• **Vertical Prediction Accuracy**
  
  – How accurate are the predicted verticals?

  \[
  V-score@l = \frac{1}{l} \sum_{r=1}^{l} \text{Accuracy}(r)
  \]
Evaluation Metric for Query Understanding

- Subtopic Diversity
  - D#-nDCG proposed by [Sakai+2011]

\[
\text{QU-score} = \lambda \text{D#-nDCG}@l + (1 - \lambda) \text{V-score}@l
\]

- Subtopic Diversity
- Vertical Prediction Accuracy

\[
\text{V-score}@l = \frac{1}{l} \sum_{r=1}^{l} \text{Accuracy}(r)
\]
Given a query and the document collection, the system is required to return a diversified ranked list documents including verticals.

<table>
<thead>
<tr>
<th>Document</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>doc-011</td>
<td>0.95</td>
</tr>
<tr>
<td>Vertical-news</td>
<td>0.80</td>
</tr>
<tr>
<td>doc-021</td>
<td>0.73</td>
</tr>
<tr>
<td>Vertical-image</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Evaluation Metric for Vertical Incorporating

- D# proposed by [Sakai+2011]

\[
D\# - nDCG = \gamma I-rec + (1 - \gamma) D-nDCG
\]

- How many intents are covered by the documents?
- How much relevant are the documents?

\[
g_i(d) = \sum_{v \in V} \delta_v(d) P(v|i) rel_i(d)
\]

Gain for intent \(i\)

Vertical type

Vertical importance

Document Relevance
We received **42 runs from 9 teams**

- **Techniques**
  - Topic Modeling
  - Syntactic Pattern
  - Word Embedding (e.g. word2vec)
  - Learning to Rank

- **Resources**
  - Web pages
  - Queries
  - Encyclopedia (e.g. Wikipedia, Baidu Baike, Hudong Baike ...)

Please join IMine-2 session at

**June 9th 11:00~12:30!**
RESULTS
NEXTI achieved the best performance – significantly different from the others.
Japanese Query Understanding subtask

Performance difference in predicting subtopic diversity.

V-score diff. > D#-nDCG diff.

Participant runs have similar performances in subtopic diversity.
ruicir achieved the best performance
– significantly different from the others
ruicir again achieved the best performance

– significantly different from the others
ruicir achieves the best performance
Researches related to INTENT/IMINE

- Many Diversification algorithms...
  - Selective Diversification [Sakai+WWW2012] [Tsukuda+AIRS2013]...
  - Hierarchical Intents [Hu+CIKM2015][Wang+SIGIR2016]...

- Evaluation Metrics [Sakai+SIGIR2011] ...

- Task-oriented Search [Yan+SIGIR2015]...

- Intrinsically Diverse Search [Raman+SIGIR2013][Umemoto+SIGIR2016]...
What would be the remaining research topics?

- **Techniques**
  - How can we generate vertical results which can satisfy different user intents?

- **Evaluation**
  - How can we evaluate the utility of a ranked list as a whole?
    - Appearance of a vertical result affects the utility of other Web documents
Summary

- **IMine-2**
  - Understand and satisfies diverse users’ search intents

- **Subtasks**
  - Query Understanding
  - Vertical Incorporating

- **Results**
  - 42 runs from 9 teams
  - Some teams achieved significantly different performance than other teams