Context-Sensitive Search
Literature Review

By: Noran Hasan
The Problem with Search as We Know It

• Search queries are often too short or too vague resulting in too many results, many of which are irrelevant.
• The large number of results are practically impossible to be fully browsed by any user.
• The use of ubiquitous devices, such as smart phones, emphasize these problems due to their limited screen space, limited ease of use, and the nature of the user’s activity while using it (usually on the move).
Typical Search Process
Possible Improvement Strategies

- Refine Query
- Modify engine
- Filter

Query → Search Engine → Results
Using Context to Improve Search

• Context can be used to improve search by one of the following approaches:
  – Using context to refine search by adding context words to query.
  – Using context to filter results and getting the most relevant.
Searching with Context

Kraft et. Al in [1] present 3 algorithms to implement contextual search:

1. Query Rewriting
   - Query Generator
   - Standard SE
   - Result set

2. Rank Biasing
   - Query Generator
   - Modified SE
   - Result set

3. IFM
   - Query Generator
   - Standard SE
   - Re-Rank
   - Result set
Searching with Context

– Query rewriting:
  • query AND context.
  • Less results (danger of low recall)

– Rank-Biasing:
  • requires a modified search engine.
  • context is an *optional* ranking term.
  • has the same level of recall as query without context.
  • E.g. <query> = <selection= cat> <optional = persian, 2>

– Iterative filtering meta-search:
  1) Candidate context words are generated
  2) Query template is applied
     – E.g. Query q and context words a b c
       » All possible combinations: qa, qb, qc, qac, qab, qac, qbc, qabc
       » If window size is 2: qab, qac, qbc
       » Force to contain high ranked term, e.g. if a is high ranked, then: qa, qac, qab, qabc
  3) Results from multiple queries are aggregated.
Desktop Context

- In desktop search, context may be inferred from:
  - Previous queries
  - URLs clicked/recently browsed documents
  - Documents on the desktop
  - Activities in standard applications
Desktop Context (cont’d)

• In [2], they rank documents not just to the relevance to current query, but to the **series of queries** issued in the **current session**.

• The table below shows some search logs:

<table>
<thead>
<tr>
<th>SID</th>
<th>Search session</th>
</tr>
</thead>
</table>
| S₁  | Ford ⇒ Toyota ⇒ GMC ⇒ Allstate
     | www.autohome.com |
| S₂  | Ford cars ⇒ Toyota cars ⇒ GMC cars ⇒ Allstate
     | www.autohome.com |
| S₃  | Ford cars ⇒ Toyota cars ⇒ Allstate
     | www.allstate.com |
| S₄  | GMC ⇒ GMC dealers
     | www.gmc.com |
Desktop Context (cont’d)

• In [3], previous queries and URLs clicked are used to identify the context.
  – If a user issues the query “Micheal Jordan” preceded by the query “NBA”, the user probably intends to query about the basketball player not the machine learning researcher at UC Berkeley.

• In [4], they summarize desktop documents and select terms to be used as context works.
Desktop Context (cont’d)

• In [5], query expansion and re-ranking are performed based on recently browsed documents with special focus on news domains.

• In [6], context is extracted from user’s activities in standard applications such as word processors. The context is used as a query to PRISM’s internal system to find a search engine that covers similar topics. If no match is found, the query is forwarded to general-purpose search engines.
How is Mobile Search Different?

• Mobile search has introduced:
  1. Limitations, in terms of limited hardware/software capabilities.
    • Not all pages can be viewed on mobile devices.
    • Deficient input emphasize the need for intelligent applications to help the user reach his target using minimum keystrokes/browsing.
  2. Increased amount of accessible of context information. Mobile devices can provide more interesting information that can be used for context.
Aiding Mobile Search

- The are several ways enhance the user’s experience while performing mobile search:
  - Query suggestions that appear as the user enters the query
  - Expanding the query or filtering search results to achieve the best precision
  - Only return results that can be viewed by the querying device
  - A combination of some/all of the above.
Dealing with Mobile H/W and S/W limitations

• In [7], they rank contents not only on the textual relevancy to the user’s query but also how adaptable it is to the target mobile environment.

• The main components in this system:
  – Delivery Context Categorization (legacy, Under DDC, DDC, Over DDC, Advanced, Computer)
  – Adaptive Context Crawling (off-line, sends multiple HTTP requests to discover if site is doing adaptation)
  – Web Page Classification (mobile optimized/mobile friendly/mobile unaware).
  – Mobile-Readiness Ranking (optimal/fair/bad/not valid)
What Kind of Context Data Do Mobile Devices Provide?

• According to [12], ubiquitous computing environments are characterized by many sensors that can sense a variety of different contexts:
  – physical contexts (like location, time)
  – environmental contexts (weather, light and sound levels)
  – informational contexts (stock quotes, sports scores)
  – personal contexts (health, mood, schedule, activity)
  – social contexts (group activity, social relationships, other people in a room)
  – application contexts (email received, websites visited)
  – system contexts (network traffic, status of printers)
Mobile Context: Location

- **Location** is one of the most important context data provided by mobile devices that desktops cannot provide.

- [8] proposes query refinement based on real-world contexts of a mobile user, such as his/her current geographic location and the typical activities at the location which are extracted by Blog mining.
Mobile Context: Location (cont’d)

• Method:
  – Mobile user inputs query
  – System Senses Real-World Contexts (obtain geographical location)
  – System Translates Contexts into Contextual Words (the typical activities at the location) by mining blog data.
  – System Assigns Weight to Contextual Word
  – Query Refinement is enforced (generate alternative queries consisting of original queries and contextual word. Let the user pick the query or return results of query with the contextual word with highest weight)

• Example: a mobile user in a bookstore issues [“da vinci”] as an original query, our system would infer from his/her original query and current place-name as a real-world context that he/she is requesting information about not a movie but a book of “da vinci code”, and offer or retrieve automatically by [“da vinci code” AND “book” AND “buy”] as one of its alternatives.

• Proposed future work: utilize not only current real-world contexts but also history of continuous past ones and/or prospective ones
Mobile Context: Location (cont’d)

• [9] proposes a system for mobile/web searches that provides the following functionality:
  – context-aware keyphrase inference
  – subtopic tree generation and context-aware re-formation
  – discovery of comparable keyphrases from the Web
  – meta vertical search focused on one subtopic

Future work:
  – utilize not only current but also past continuous and/or prospective real-world contexts
  – personalize context-based weightings based on user profiles and/or personal query logs
Mobile Context: Location (cont’d)

- [10] presents a new search browser interface in which context, in the form of time and location information, is integrated with preference information derived from the queries of like-minded communities of mobile searchers.
- The sample communities the prototype used are entertainment and tourist sites.
Mobile Context: Location (cont’d)

• In [11], mobile phone users are encouraged to record multimedia blogs on-the-fly, enriched with inputs from other physical sensors.
• The blogs are geotagged and uploaded to a remote server that positions these blogs on a spatial platform (e.g. map).
• Internet users can zoom into any part of the map and browse multimedia blogs at those locations. Moreover, users may query selected regions for desired information.
• The Internet user selects a region and sends a query to phones in that region.
• The phone user replies to it, and the reply is transmitted back to the server.
• The query and reply are associated to the blog.
References


2) Towards Context-Aware Search by Learning A Very Large Variable Length Hidden Markov Model from Search Logs © 2009

3) Context-Aware Query Classification © 2009

4) Summarizing Local Context to Personalize Global Web Search ©2006

5) Using the Current Browsing Context to Improve Search Relevance © 2008

6) Towards Context-Based Search Engine Selection ©2001

7) A crawling and ranking method for improving context-awareness in mobile web search © 2008
References

8) Activity-Based Query Refinement for Context-Aware Information Retrieval © 2006
9) ReCQ: Real-World Context-Aware Querying © 2007
13) Query Suggestions for Mobile Search: Understanding Usage Patterns ©2008