Providing Responses to Local-intent Queries from Web Results

N/A
Providing Responses to Local-intent Queries from Web Results

ABSTRACT

Local search engines use stored structured data to answer user queries relating to businesses and other places of interest in specific locations. When such structured data is unavailable, a local search engine cannot provide useful responses to user queries. This disclosure describes techniques to provide responses to local-intent queries based on snippets from web results. Responses are provided that include attributes such as phone number, address, etc. of local businesses, and businesses within specified categories (e.g., restaurants, repair shops, etc.).

KEYWORDS

- Local-intent query
- Web search
- Local search
- Attribute query
- Categorical query
- Response passage

BACKGROUND

A local search engine (LSE) aims to answer queries for points of interest (e.g., businesses) near the current location of the user. For an LSE to work, the relevant points of interest need to be represented in the structured data of the engine. Queries that seek the attributes of a local business, entity, or place of interest, e.g., phone number, email address, physical address, business hours, etc., are known as attribute queries. If a business is not known to the LSE, attribute queries relating to the business cannot be answered by the LSE.
Queries that seek businesses by their category, e.g., “air-conditioning service near me,” “restaurant near me,” etc., are known as categorical queries. When possible, such queries are answered by an LSE using, e.g., a response package of a small number of businesses in the desired category. However, a business that is not represented in the LSE database does not appear in responses to categorical queries.

Besides, there can be categorical queries for which an LSE cannot produce relevant responses, even if the corresponding businesses are present in the LSE database. This occurs when the query-to-place connection is not immediately clear. An example of such a query is “Aadhaar processing center near me.” In this query, “Aadhaar” refers to a unique ID card issued by the Government of India; applications for Aadhaar can be processed at many banks. Although the locations of banks are indexed and available to an LSE, the inference that banks (e.g., specific locations of bank branches) are a relevant response to an Aadhaar-related user query is not possible. Web retrieval and ranking, however, implicitly can make the necessary inference, as the top-ranked search results, from which responses are generated, include lists of relevant places.

Search engines sometimes provide results including snippets that are extracted from web documents. Queries that are answer-seeking, e.g., “why is the sky blue,” can then be used to pose attribute queries, e.g., “what is the phone number of sunglasses-for-less.” However, snippet-style responses have certain shortcomings for providing responses to attribute queries such as:

- Snippet-style responses do not specifically comprise passages that include a business name and attribute values. Such passages may be generated by existing passage builders, but this is rare.
• For scoring candidate responses, the presence of an attribute matching the query intent in the passage is not considered. For example, when the query is for an email address, the presence of an email address in the response does not increase the score. Instead, scoring relies on textual tokens, e.g., labels to match the query tokens. For example, for a query “cascale email,” (cascale being the name of a business) scoring would only boost the score of a response if the token “email” appeared in the response, but not if there is an email address actually present. This also applies to other attributes such as physical addresses, phone numbers, and business hours.

Snippet-style responses have shortcomings for categorical queries, including:

• Snippet-style responses generally do not process categorical queries in a manner that is expected by users, e.g., do not take into account that such queries likely seek lists of businesses.

• Snippet-style responses do not account for geographic distance. Users typically query for businesses near their location or near an explicit location in the query. Snippet-style responses may be scored higher if they are in the same locality, e.g., if locality names are mentioned in both query and response, but do not specifically account for physical distance.

DESCRIPTION

This disclosure describes techniques to process attribute queries and categorical queries in a manner that provides responses to local-intent queries from web results.

Attribute queries

![Diagram of processing an attribute query to obtain local-intent results]

Fig. 1: Processing an attribute query to obtain local-intent results
As illustrated in Fig. 1, the processing of an attribute query to obtain local-intent results comprises the following components.

- **Pre-trigger model** (102): A pretrigger model identifies business attribute queries, e.g., queries for business address, phone number, email address, opening hours, etc.

- **Query-intent determiner** (104): A query-intent determiner uses the following sources to determine query intent:
  1. Parsing from a “local intent grammar” which is a grammar for local-intent queries;
  2. Salient terms for local intent, e.g., terms aggregated over historical queries for which there is a local-intent grammar parse;
  3. As a fallback, trigger tokens e.g., “email,” for intents not well modeled by either grammar or salient terms.

- **Passage builder** (106): A passage builder generates response passages that include business names and values for the attribute matching the query intent. Passages are built from components such as value annotations (e.g., phone number); the nearest heading, which is often the business name; the sentence between the heading and the value, which often provides additional content, e.g., department; etc. Good passages include value annotations that match query intent and enough context for scoring.

- **Scorer** (108): A scorer generates scores for possible responses. The scoring function used by the scorer boosts scores if the response contains an attribute value matching the query intent. The boost depends on the attribute type, e.g., email, address, open hours, etc., such that responses with different attributes can get boosted by different amounts to account for different per-attribute precision/recall tradeoffs.
Categorical queries

As illustrated in Fig. 2, the processing of an attribute query to obtain local-intent results comprises the following components.

- **Pre-trigger model** (202): A pre-trigger model identifies queries that are categorical.
- **Text-passage generator** (204): A text-passage generator generates text passages containing essential information relating to a single business. A text-passage can be based on local annotations (to obtain place data for local businesses); headings (to construct passages comprising the closest previous heading and the last local annotation before the next heading); patterns in pages; etc.
- **Response package builder** (206): A response-package builder assembles \( n \)-packs of the business passages generated by the text-passage generator. For example, if \( n=3 \), then passages relating to three businesses are assembled by the response package builder. The response package builder also finds a common heading, which can be the category name.
- **Scorer** (208): A scorer scores the \( n \)-packs of business passages based on the geographical distance between the businesses therein and either the query location or, with user permission, the user location (if there is no explicit location in the query, or if the location in the query is of the form “near me”). For response scoring based on the distance to the user or the query, the geo-location for each candidate passage is determined from the latitude/longitude (lat-long) geocoding of any address contained within the response, or the lat-long geocoding of any location entity within the response, or from the location of the document from which the response was extracted.

Fig. 2: Processing a categorical query to obtain local-intent results
Further to the descriptions above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information (e.g., information about a user’s social network, social actions or activities, profession, a user’s preferences, or a user’s current location), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user’s identity may be treated so that no personally identifiable information can be determined for the user, or a user’s geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes techniques to provide responses to local-intent queries based on snippets from web results. Responses are provided that include attributes such as phone number, address, etc. of local businesses, and businesses within specified categories (e.g., restaurants, repair shops, etc.).