SEQ: Example-based Query for Spatial Objects

Siqiang Luo, Jiafeng Hu, Reynold Cheng, Jing Yan, Ben Kao

The University of Hong Kong

Motivation

Question: when we are searching multiple spatial objects with complex requirements, what are we going to do?

Option 1: SQL?

Difficult for non-expert users

Option 2: Keyword search?

Hard to describe spatial layout characteristics

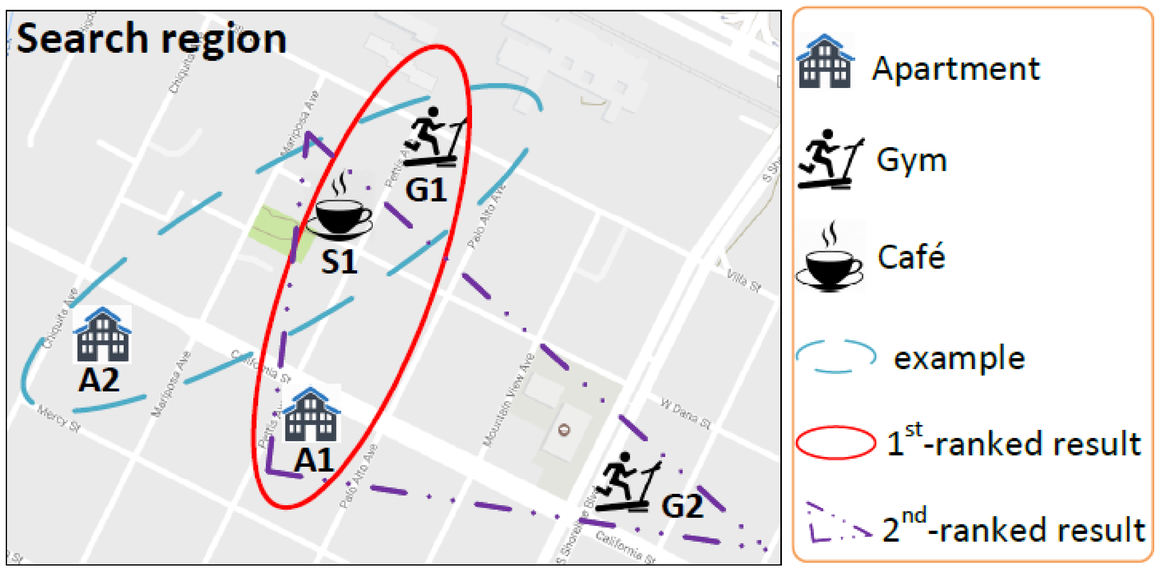
Hence, we are promoting more user-friendly search: **example-based spatial search**

SEQ: Spatial Exemplar Query

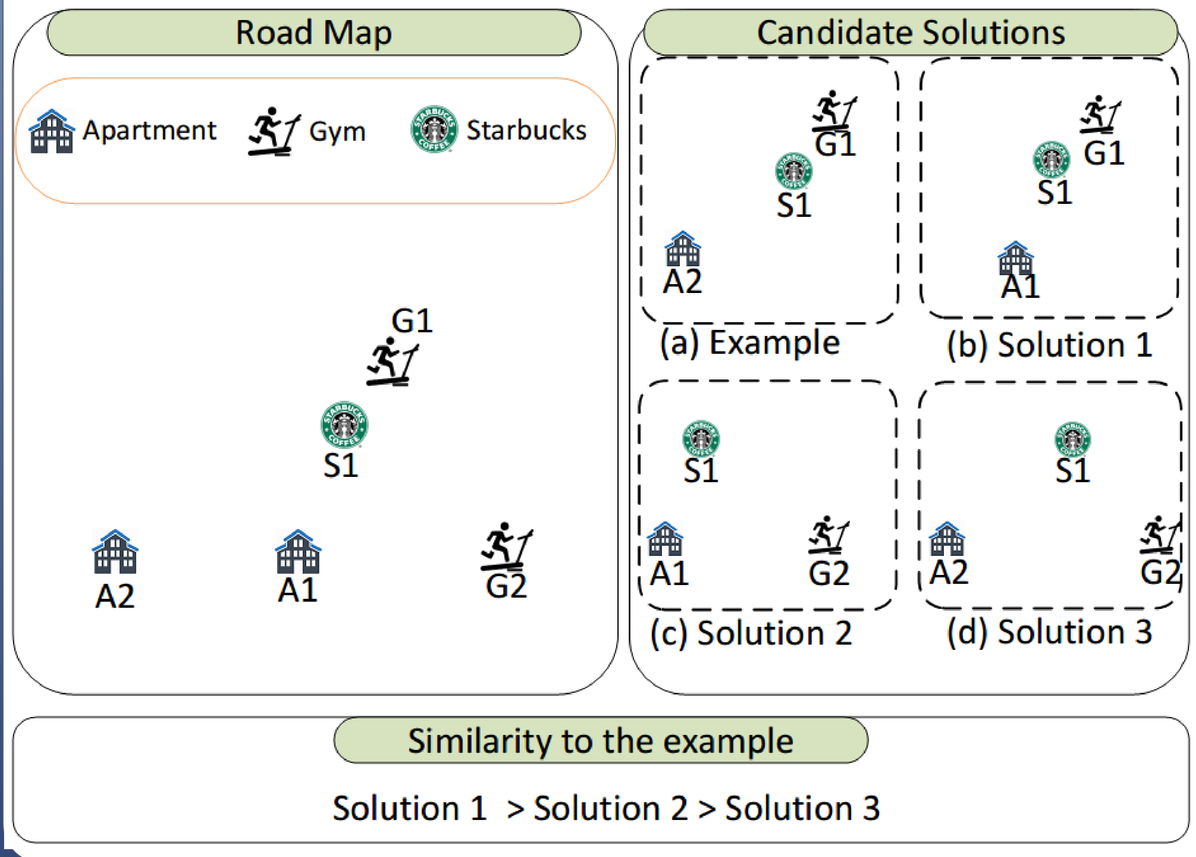
Consider renting a house:
Find an apartment which is close to a gym, and there is a café between them.

Input: mark an apartment, a gym and a café as a desired example
Output: A list of similar results

Interface view



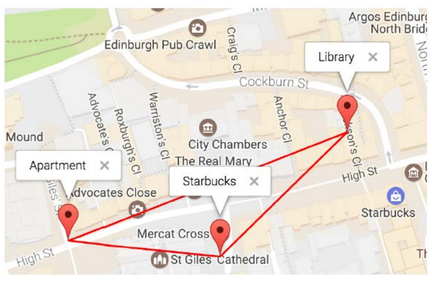
Analytical view



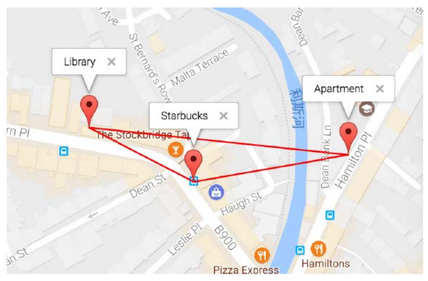
Query Definition

Definition (SEQ) Given a spatial range R, an integer k, an example tuple t. The SEQ returns top-k similar tuples t_1, t_2, …, t_k with respect to the tuple similarity to t, such that all objects in t_i are located in R, and t_i is of the same category as t.

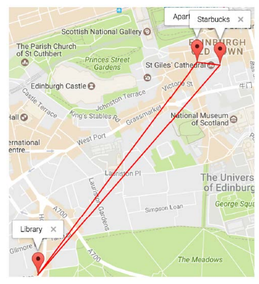
Tuple Similarity: $\alpha \times$ Spatial Similarity + $(1-\alpha) \times$ Attribute Similarity



User input example



Attribute ✓
Spatial ✓



Attribute ✓
Spatial ✗

Algorithms

1. Find all candidates, e.g., find all (Apartment, Gym, Café)
2. Object-wise ranking based on attribute similarity
3. Depth first search and prefix-based pruning

Given example t=(a, b, c) and candidate t_1=(d, e, f), we can compute a lower bound of Similarity(t, t_1) based on (a,b) and (d,e). If the lower bound is larger than k-th similarity, then prune t_1.

Experiments

Yelp dataset (https://www.yelp.com/dataset_challenge)

POI information: location, category, rating, review count

α: weight of spatial similarity **k**: #returned results
r: the radius of search region (km)
Pruning effectiveness: the percentage of pruned candidates by prefix-based pruning

α	0.1	0.3	0.5	0.7	0.9
Time (ms)	316	336	338	349	380
Pruning effectiveness (%)	98.8	98.7	98.5	98.1	97.1
k	1	5	10	20	50
Time (ms)	32	338	349	363	403
Pruning effectiveness (%)	99.8	98.5	98.1	97.5	96.5
r	1	2	3	4	5
Time (ms)	68	161	338	586	847
Pruning effectiveness (%)	97.6	98.3	98.5	99.5	99.5

Future Work

1. Integrate SEQ into real spatial services
2. human-in-the-loop SEQ
3. Different ways of inputting examples

References

[1] G. Cong, C. S. Jensen, and D. Wu. Ecient retrieval of the top-k most relevant spatial web objects. VLDB, 2(1):337–348, 2009.
[2] H. Li, C. Chan, and D. Maier. .ery from examples: An iterative, data-driven approach to query construction. VLDB, 8(13):2158–2169, 2015.
[3] J. Liu, K. Deng, H. Sun, Y. Ge, X. Zhou, and C. S. Jensen. Clue-based spatio-textual query. VLDB, 10(5):529–540, 2017.