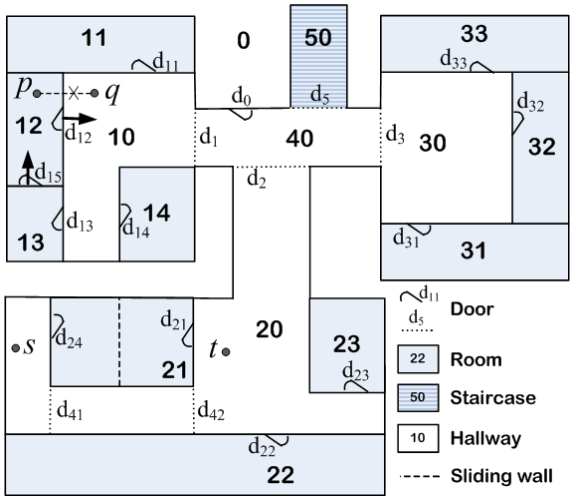
Efficient Distance-Aware Query Evaluation on Indoor Moving Objects

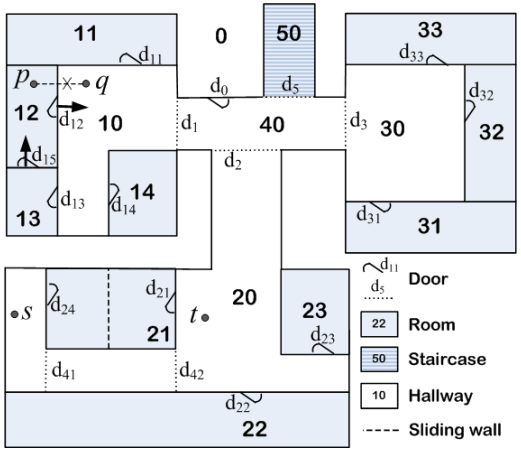
Scott Xike XIE

Indoor Space

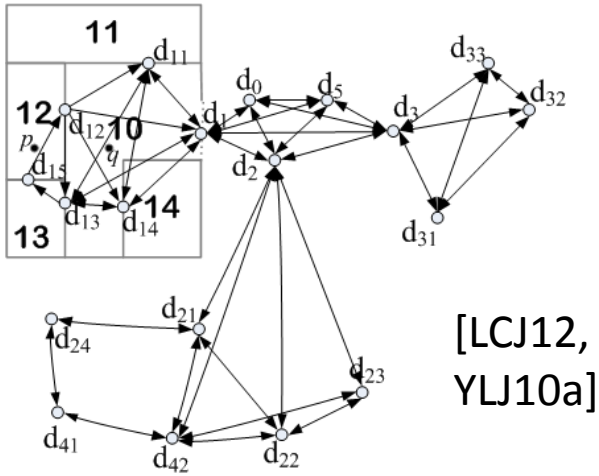
- Indoor Positioning Technologies
  - Assisted GPS (A-GPS), Wi-Fi, RFID and Bluetooth
- Indoor Distance-aware Location-based Service
  - Nearby service notification, advertisement
- Indoor Distance characteristics
  - Topological constraints
  - Temporal variations
  - Locational uncertainties



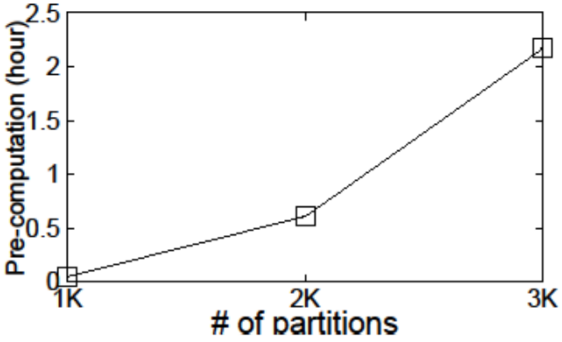
Topological Constrains and Temporal Variations



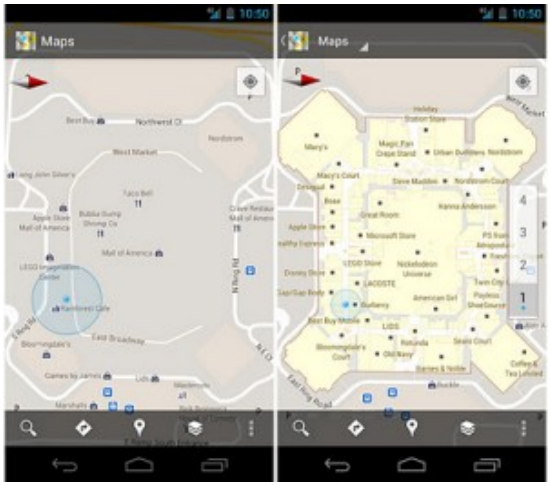
$|q,p| = |q \rightarrow p|$

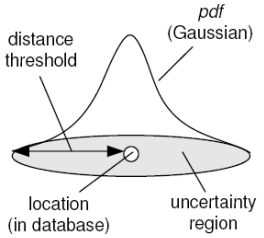


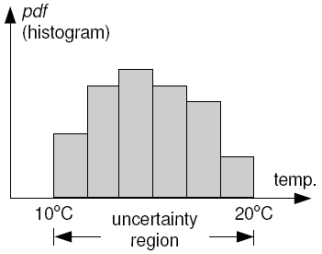
[LCJ12, YLJ10a]



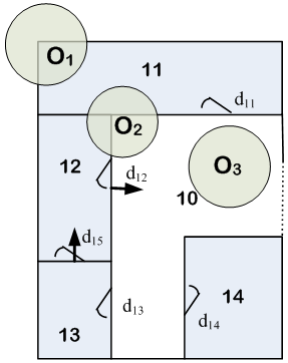
Location Uncertainty

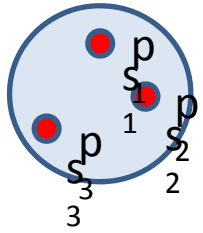






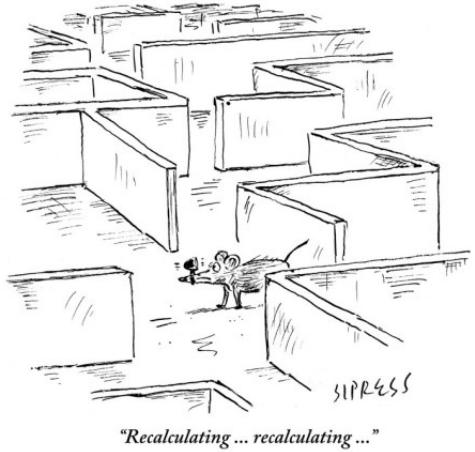
[Wolfson98, Pfoser99, Deshpande04, Agrawal00, Agrawal08, Cheng03, LC08, LP07]





$$O = \{(s_i, p_i)\} \sum_{s_i \in O} p_i = 1$$

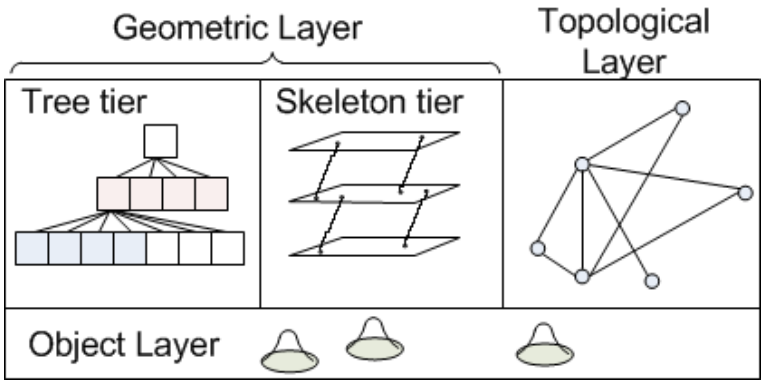
$$O = \cup_{1 \le j \le m} S[j]$$



"Recalculating ... recalculating ..."

Overview

- Indoor distances
  - (door-door distance, door-point distance, Expected Indoor distance)
- Composite Indoor Index (for Indoor space and objects)
- Distance-aware queries (range query, k-NN)

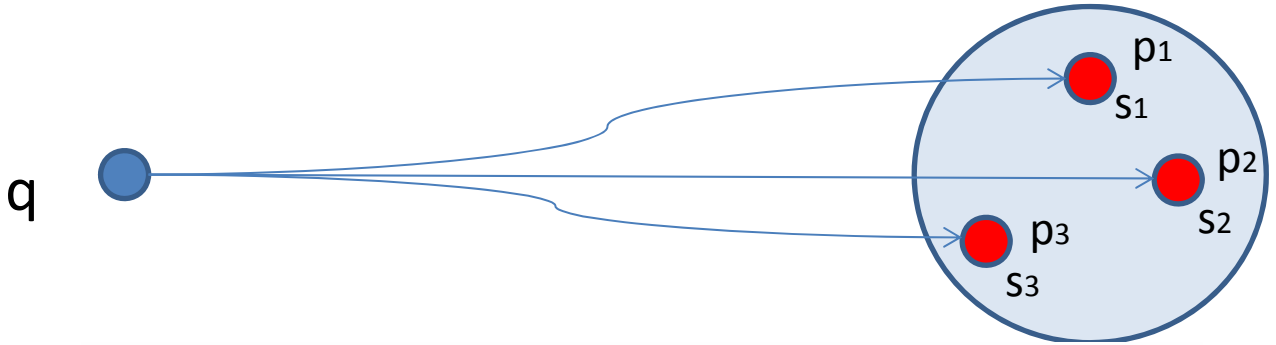


Existing works	Topological Constrains	Temporal Variations	Locational Uncertainties
[LCJ12]	√	×	×
[YLJ10a]	√	×	√
[YLJ10b]	√	×	√
This work	√	√	√

Agenda

- Introduction
- Indoor distances
  - Expected indoor distances
  - Indoor distance categories
  - Upper/lower bounds
- An composite indoor index
  - Three layers, two tiers, and
- Indoor distance-aware queries
  - Entity-based query (range query)
  - Rank-based query (k nearest neighbor query)
- Results
- Conclusions

Expected Indoor Distances



$$|q, O|_I = E_{s_i \in O}(|q, s_i|_I) = \sum_{s_i \in O} |q, s_i|_I \cdot p_i$$

Pros

- Interpretative in entity-based query (range query)
- Semantically comprehensive in rank-based query (kNN query) **[CLY09]**

Cons

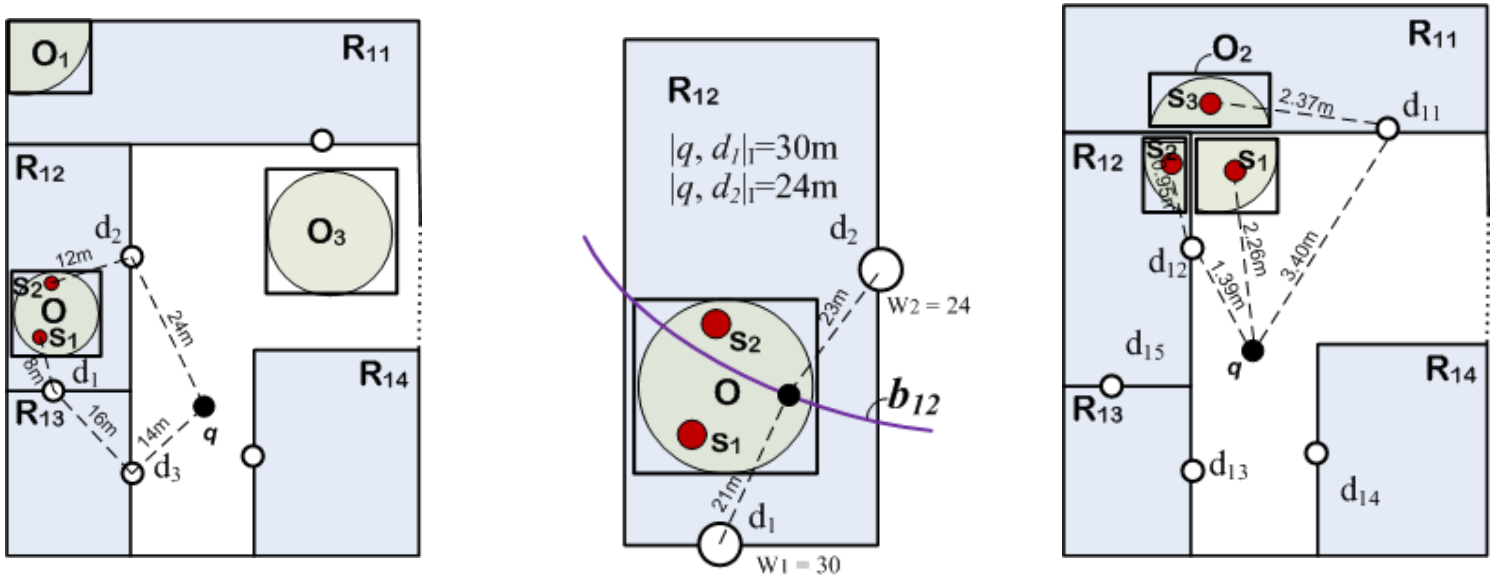
- Expensive to calculate

How to solve?

To use upper/lower bound

Cases of Indoor Distances

- Single-partition Single-path Distance
- Single-partition Multi-path Distance
- Multi-partition Path Distance

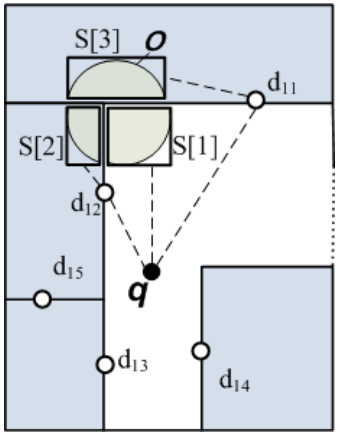


Note: to evaluate |q, O|I, we have find shortest paths from q to all O's instances.

Upper/lower Bounds for Indoor Distances

- Euclidean Lower Bound
- Topological Upper/lower Bound
  - Topological Lower Bound $[|q, O|_{minI}, |q, O|_{maxI}]$
  - Topological Upper Bound
- Probabilistic Upper/lower Bound
  - Markov Lower Bound
  - Probabilistic Upper/lower Bound

Probabilistic Upper/lower Bound

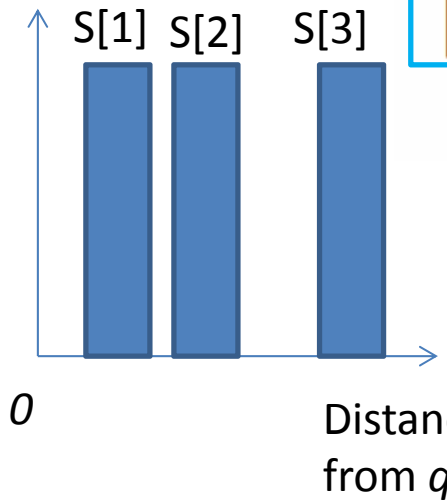


Lemma 4: **(Markov Lower Bound)**

$$E(|q,O|_I) \geq |q,S[i]|_{maxI} \cdot (1-\hat{p}_i)$$

Lemma 5: **(Probabilistic ULBounds)**

$$|q,S[i]|_{maxI} \cdot (1-\hat{p}_i) + |q,O|_{minI} \cdot \hat{p}_i \leq E(|q,O|_I) \leq |q,O|_{maxI} \cdot (1-\hat{p}_i) + |q,S[i]|_{maxI} \cdot \hat{p}_i$$



Summary of Indoor Distances

INDOOR DISTANCES AND THEIR UPPER / LOWER BOUNDS

Indoor Distance	Bounds
Single-partition single-path distance Single-partition multi-path distance	Indoor Topological Upper/ Lower Bounds (Equation 7)
Multi-partition path distance	Indoor Probabilistic Upper/ Lower Bounds (Equation 8)

Pros

- Indoor distance could be well-approximated, thus computational efforts are saved

Cons

- Require calculating shortest path from q to an object.
- Expensive for a large graph.

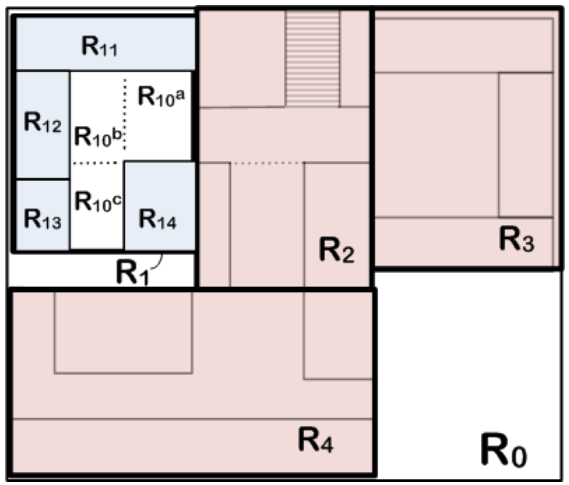
How to solve?

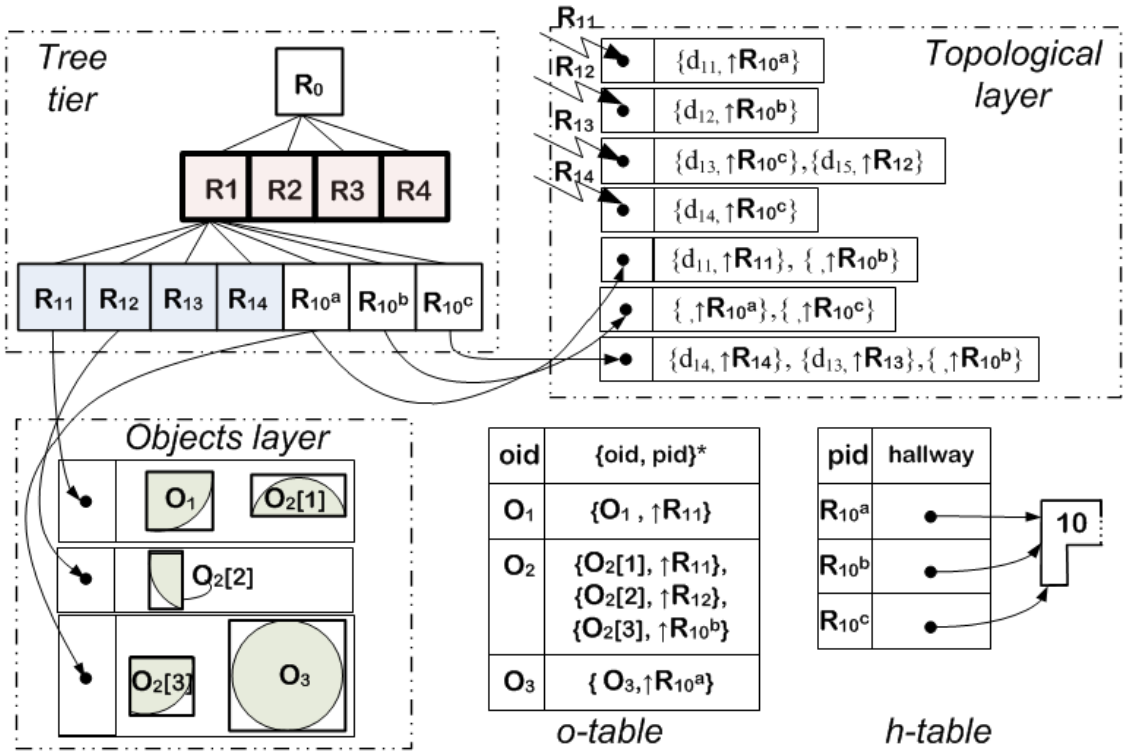
To retrieve a sub-graph by using Pruning techniques.

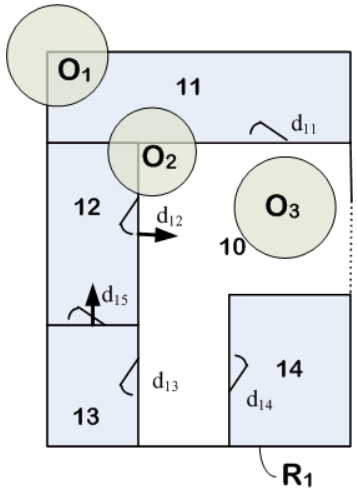
Agenda

- Introduction
- Indoor distances
  - Expected indoor distances
  - Indoor categories
  - Upper/lower bounds
- An composite indoor index
  - Three layers, two tiers, and auxiliary structures
- Indoor distance-aware queries
  - Entity-based query (range query)
  - Rank-based query (k nearest neighbor query)
- Results
- Conclusions

COMPOSITE INDEX FOR INDOOR SPACES



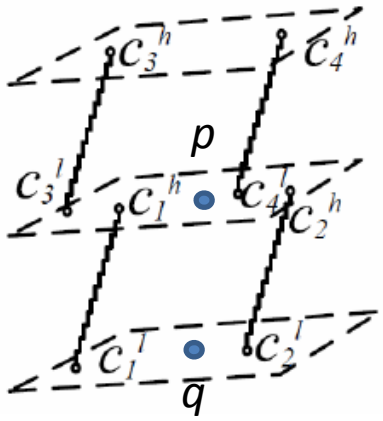


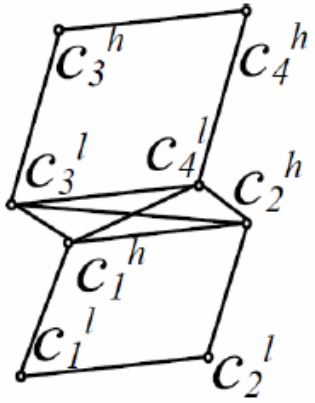


(without the skeleton tier)

COMPOSITE INDEX FOR INDOOR SPACES

- Tree tier
  - branch-and-bound
- Skeleton tier
  - Skeleton distance
  - Geometric Lower Bound Property
- Dynamic Operation
  - Topological Layer Operations
  - Object Layer Operations

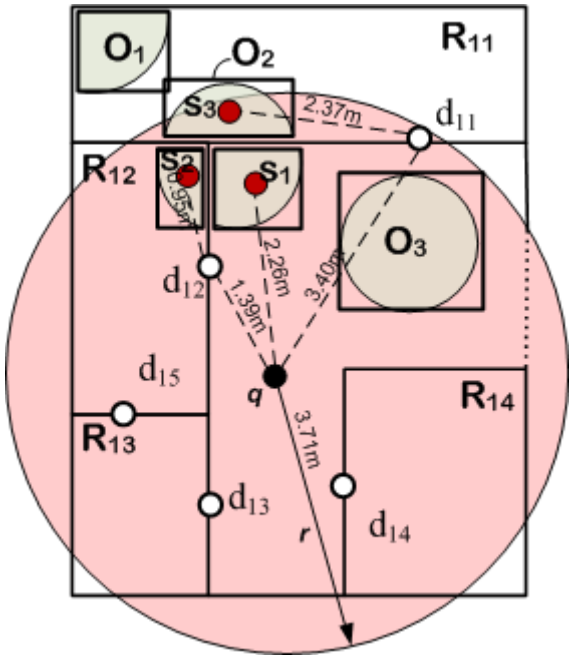




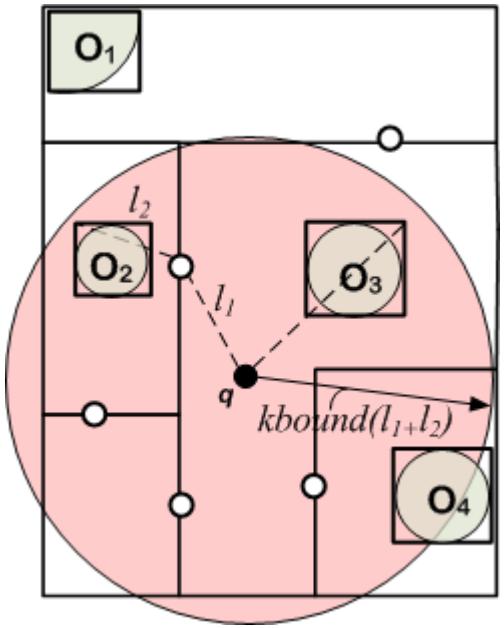
Agenda

- Introduction
- Indoor distances
  - Expected indoor distances
  - Indoor distance categories
  - Upper/lower bounds
- An composite indoor index
  - Three layers, two tiers, and
- Indoor distance-aware queries
  - Entity-based query (range query)
  - Rank-based query (k nearest neighbor query)
- Results
- Conclusions

Queries



*i*RQ



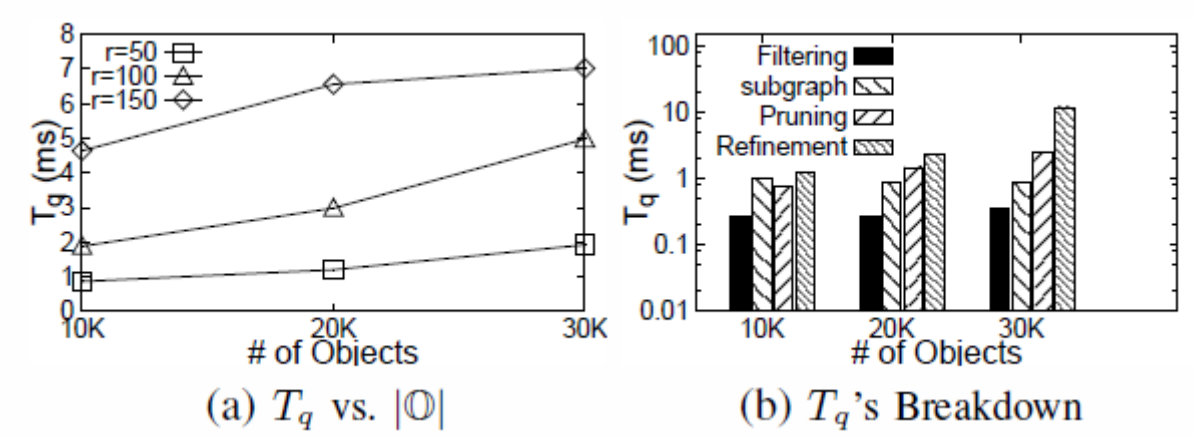
*i*kNNQ

Experimental Studies

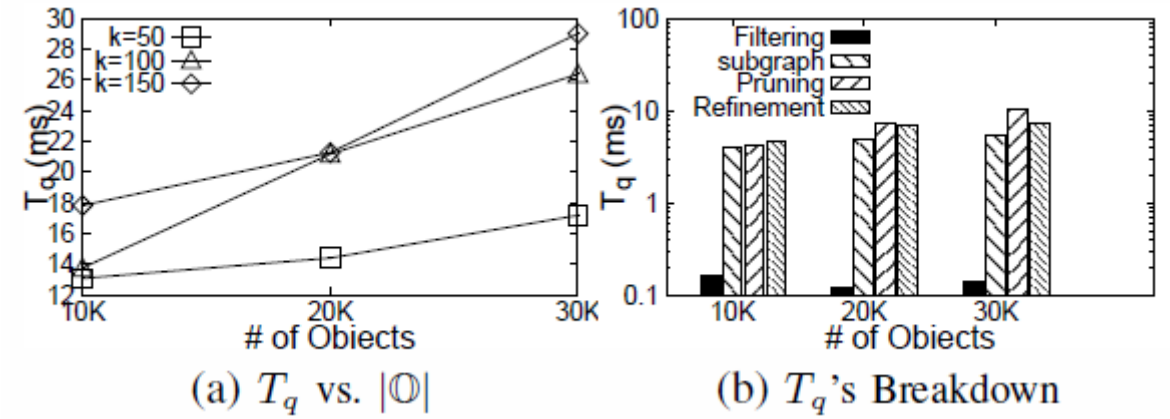
	Parameters	Values
Indoor Space	Size of each floor	600m*600m*4m
	# of rooms per floor	100 rooms
	# of floors	10, **20**, 30 floors
Indoor Moving Objects	# of objects	10k, **20k**, 30k
	Uncertainty	5, **10**, 15 meters
	Probability distribution	Gaussian Distribution
	# of instances per object	100
Queries	iRQ	50, **100**, 150 m (range)
	ikNNQ	50, **100**, 150 (k)

*i*RQ & *i*kNNQ

*i*RQ

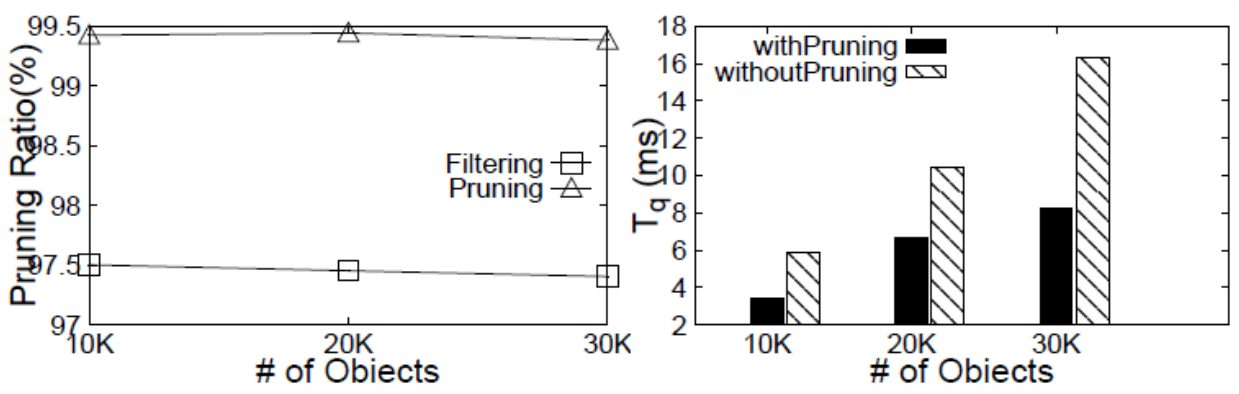


*i*kNNQ



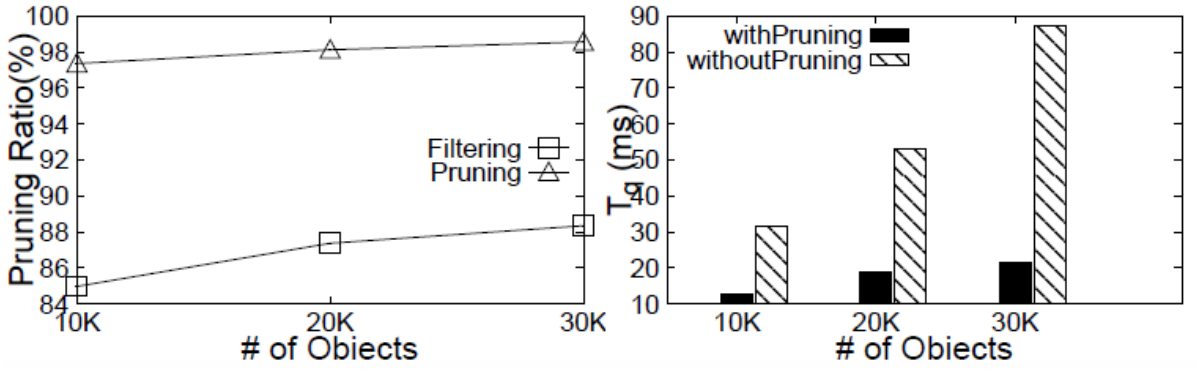
Effectiveness of Indoor Distance Bounds

*i*RQ



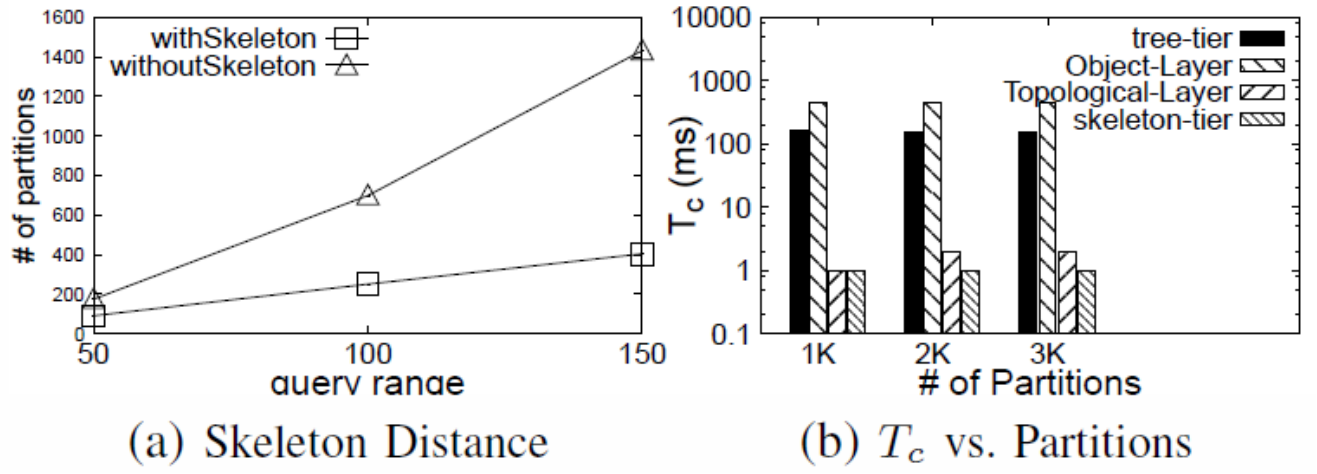
(a) Filtering & Pruning (*i*RQ) (b) Pruning Phase (*i*RQ)

*i*kNNQ



(c) Filtering & Pruning (*i*kNNQ) (d) Pruning Phase (*i*kNNQ)

Analysis



(a) Skeleton Distance (b) T_c vs. Partitions

Conclusion

- We investigate indoor distances w.r.t. indoor uncertainties and topologies;
- For efficiency, we derive upper/lower bounds for indoor distances;
- We design a composite index to facilitate indoor queries;
- Experiments show the efficiency, effectiveness, and scalability.

- Future work
  - Study other query semantics;
  - Study how to reuse the topological computation efforts for multiple, related queries.

Q & A

Thanks!

Reference

- [LCJ12] H. Lu, X. Cao, and C. Jensen. A foundation for efficient indoor distance-aware query processing. In ICDE, 2012.
- [YLJ10a] B. Yang, H. Lu, and C. Jensen. Probabilistic threshold k nearest neighbor queries over moving objects in symbolic indoor space. In EDBT, 2010.
- [YLJ10b] B. Yang, H. Lu, and C. Jensen. Scalable continuous range monitoring of moving objects in symbolic indoor space. In CIKM, 2009.
- [Wolfson98] A. Prasad Sistla, Ouri Wolfson, Sam Chamberlain, and Son Dao. Querying the uncertain position of moving objects. In Temporal Databases, 1998.
- [Pfoser99] D.Pfoser and C. Jensen. Capturing the uncertainty of moving-objects representations. In SSDBM, 1999.
- [Deshpande04] A. Deshpande, C. Guestrin, S. Madden, J. Hellerstein, and W. Hong. Model-driven data acquisition in sensor networks. In VLDB, 2004.
- [Agrawal00] Rakesh Agrawal and Ramakrishnan Srikant. Privacy-preserving data mining. In SIGMOD, 2000.
- [Aggarwal08] Charu C. Aggarwal. On unifying privacy and uncertain data models. In ICDE, 2008.
- [Cheng03] R. Cheng, D. Kalashnikov, and S. Prabhakar. Evaluating probabilistic queries over imprecise data. In SIGMOD, 2003.
- [LC08]X. Lian and L. Chen. Monochromatic and bichromatic reverse skyline search over uncertain databases. In SIGMOD, 2008.
- [HP07]H.-P. Kriegel, P. Kunath, and M. Renz. Probabilistic nearest-neighbor query on uncertain objects. DASFAA, 2007
- [CLY09] Cormode, F.Li, and K. Yi. Semantics of ranking queries for probabilistic data and expected distance, ICDE 2009