spDB presentation

SCUBA: Scalable Cluster-Based Algorithm for Evaluating Continuous Spatio-Temporal Queries on Moving Objects

(Advances in Database Technology - EDBT 2006)

paper by

Rimma V. Nehme and Elke A. Rundensteiner

presentation by

Rune Leth Wejdling

November 2008

Outline

- The idea and motivation
- SCUBA algorithm
- Optimisation
- Results
- Conclusion
- Related work and contributions
- Evaluation of paper

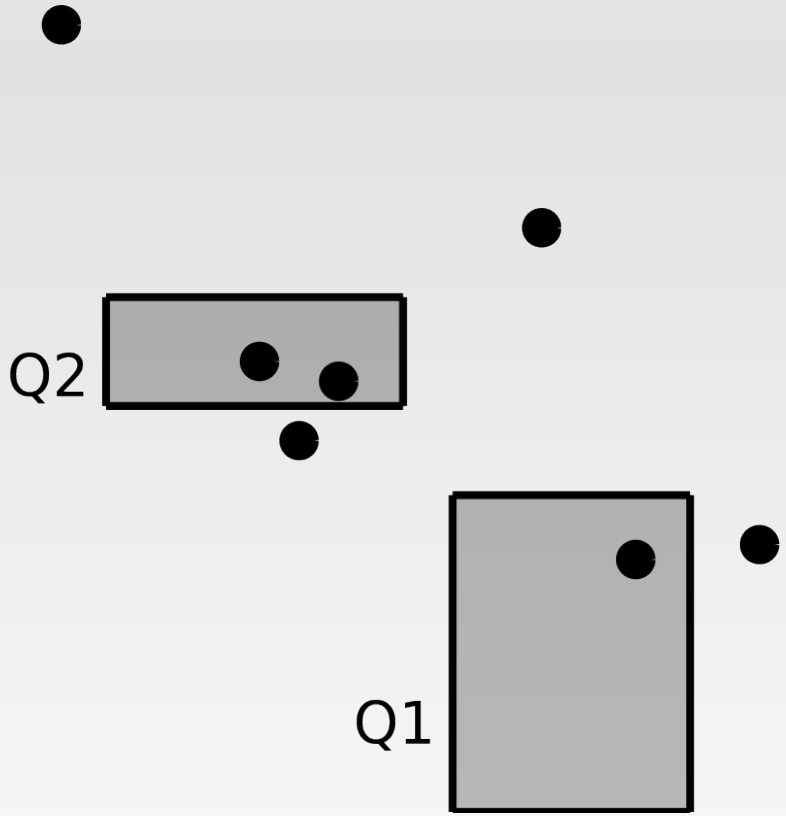
The idea and motivation (1/2)

- Many GPS enabled devices enables LBS.
- Many LBSs rely on continuous spatio-temporal queries.
  - “How many cars are at my favourite gas station right now”?
- Many moving objects naturally move in groups.
  - Cars on a highway, people at rock concerts, migrating birds etc.

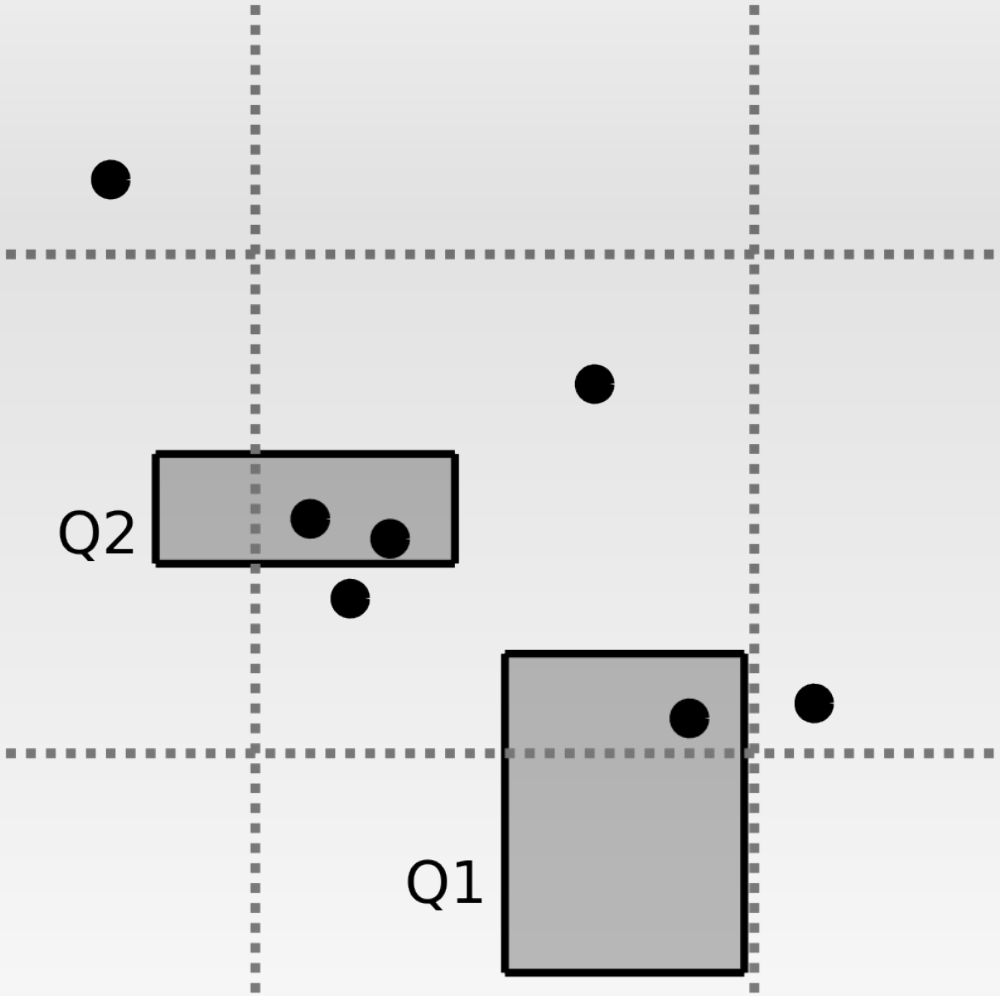
The idea and motivation (2/2)

- Groups moving in the same direction with the same speed can be utilised to create *moving clusters.*
- When clustering objects and queries with similar spatial-temporal attributes, many joins between *continuous queries* and *moving objects* can be simplified.

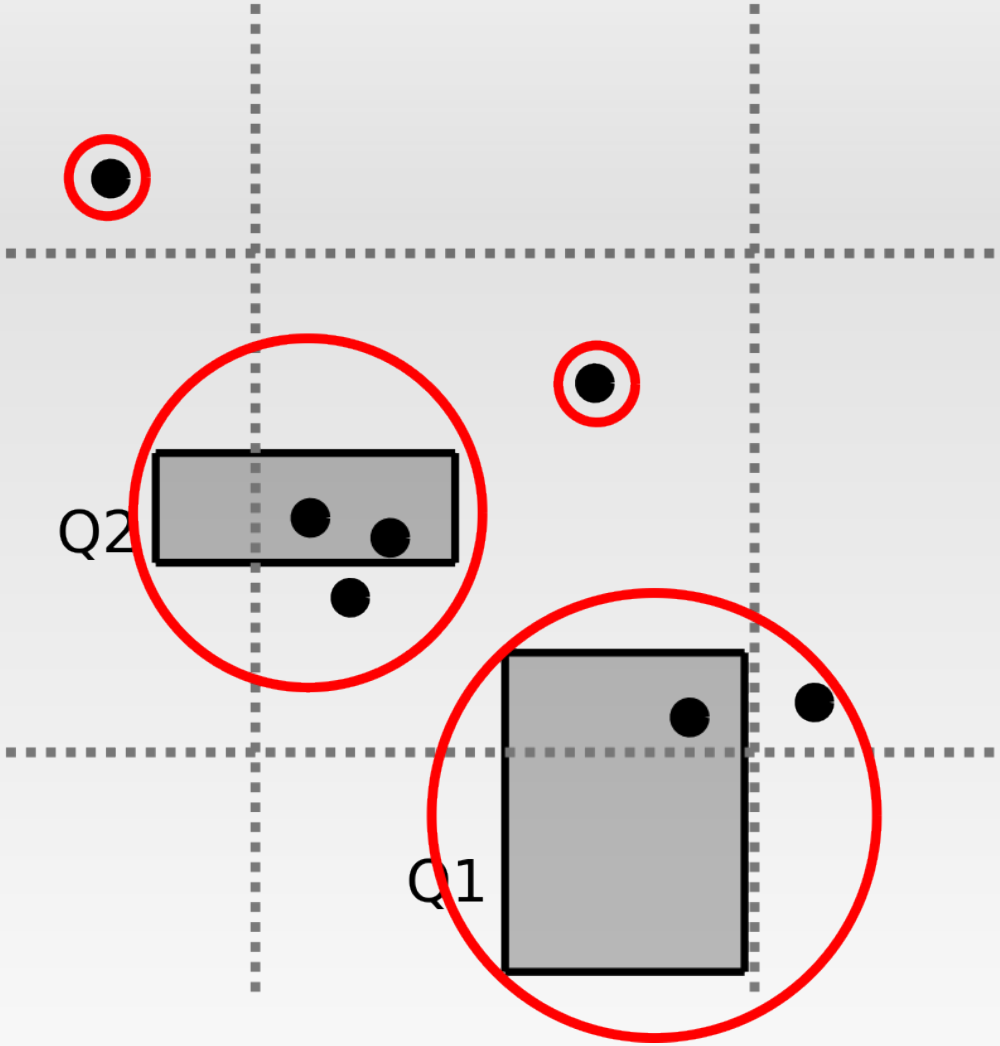
Moving objects and queries



Grid based index



Grid cluster based index



SCUBA overview

- Motion Model
  - Moving objects, Continuous queries and Moving clusters
- Join algorithms
  - *Join-between* and *join-within*
- *Scalable Cluster-Based Algorithm* (SCUBA)
  - Putting it all together

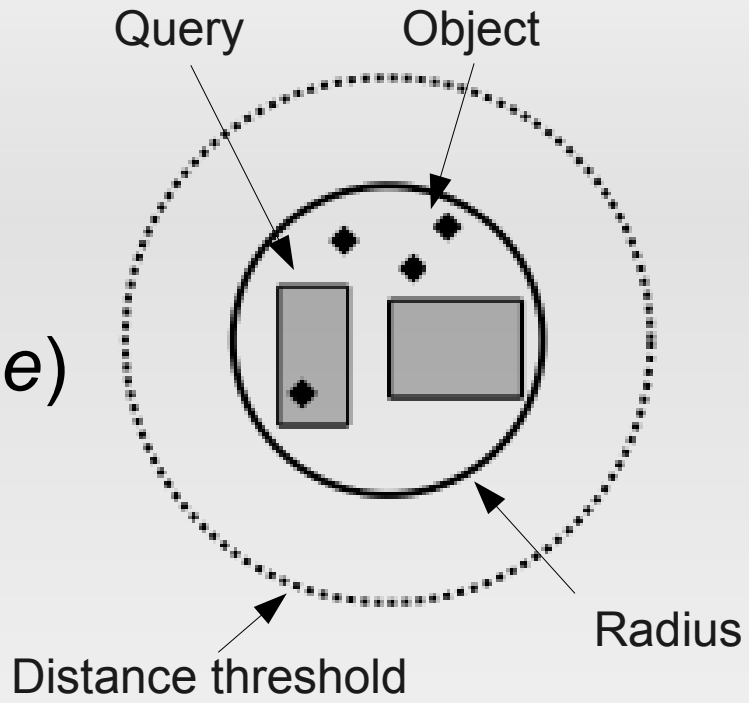
The Motion Model (1/2)

- Moving object
  - (*OID*, Loc_t, *t*, *Speed*, *CNLoc*, *Attrs*)

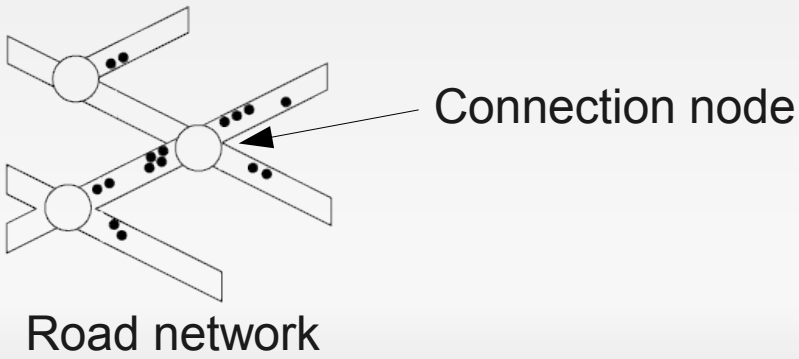
- Continuous query
  - (*QID*, Loc_t, *t*, *Speed*, *CNLoc*, *Attrs*)

The Motion Model (2/2)

- Moving cluster
  - (*CID*, *Loc*$_t$, *n*, *OIDs*, *QIDs*, *AveSpeed*, *CNLoc*, *R*, *ExpTime*)



- Connection nodes
  - E.g. intersections in a road map with high likelihood of splitting clusters

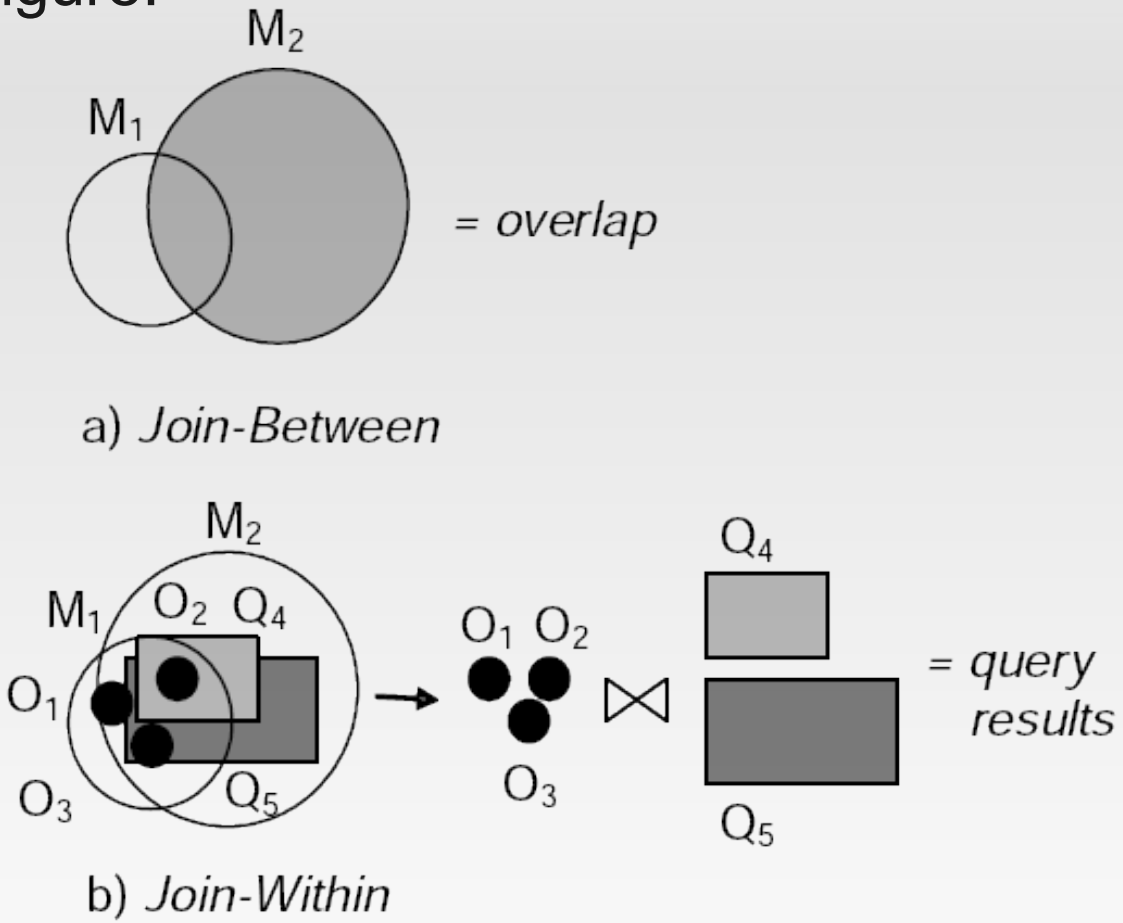


Moving cluster formation

- Location update procedure for a object:
  1. Locate grid cell for location
  2. Find clusters in grid cell
  3. If empty create new cluster
  4. Otherwise find cluster with similar spatio-temporal attributes
  5. If cluster found, join, else create new cluster

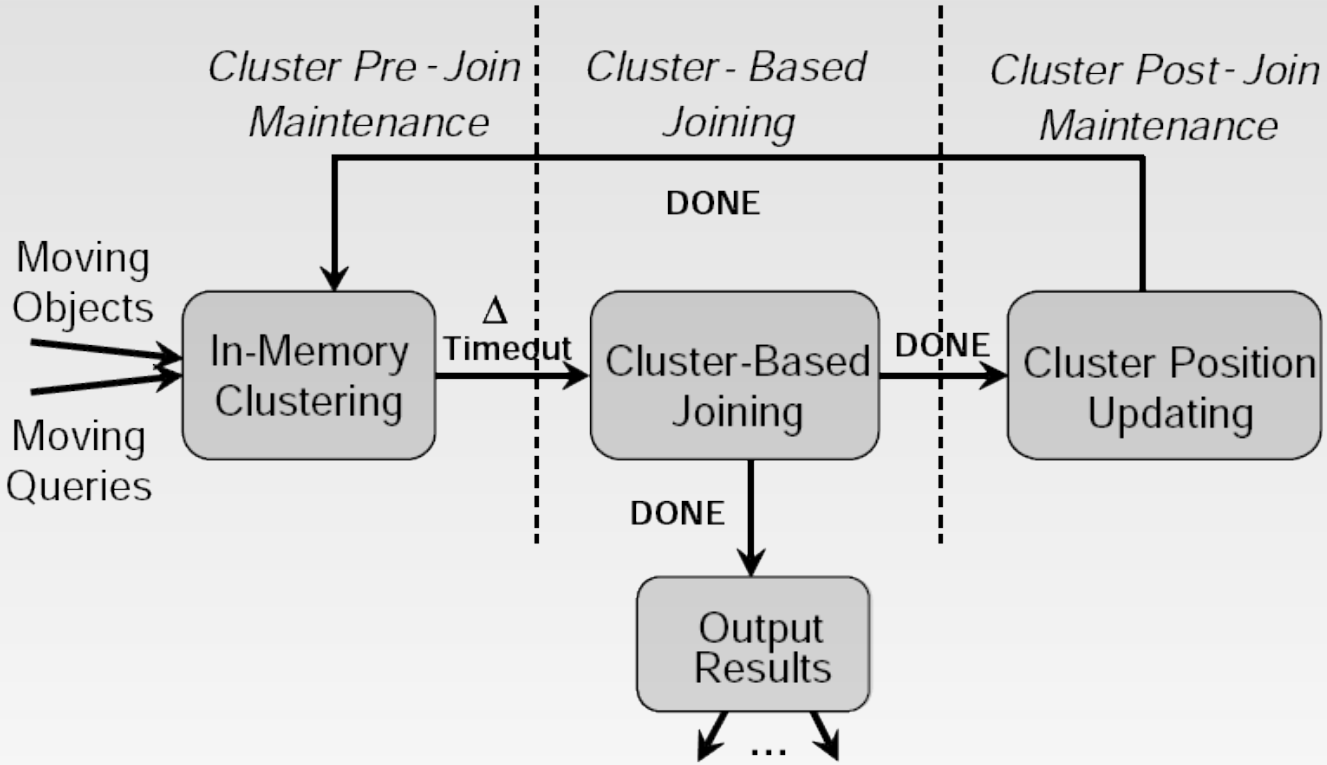
Join algorithms

Best explained by figure:

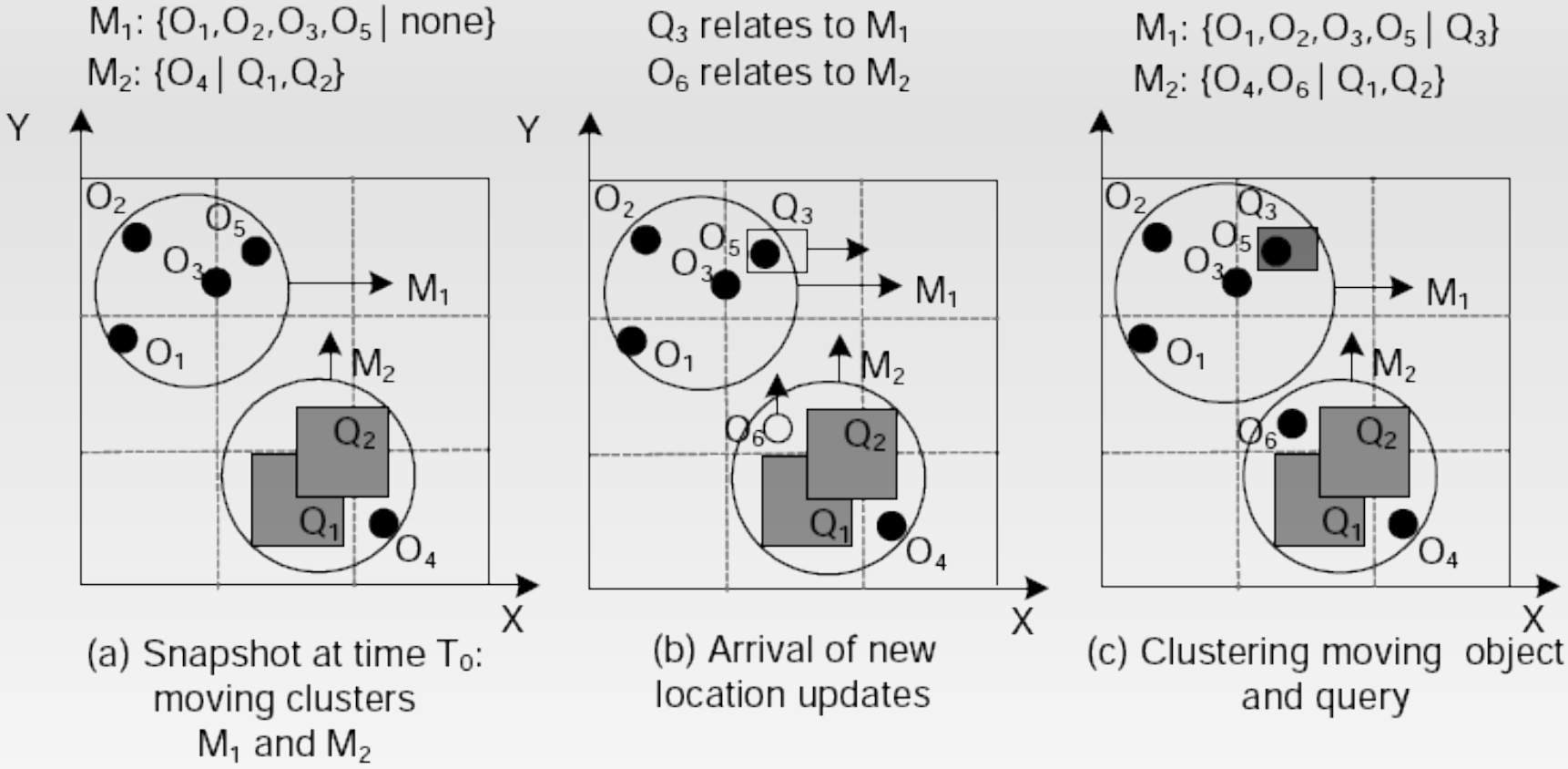


SCUBA algorithm

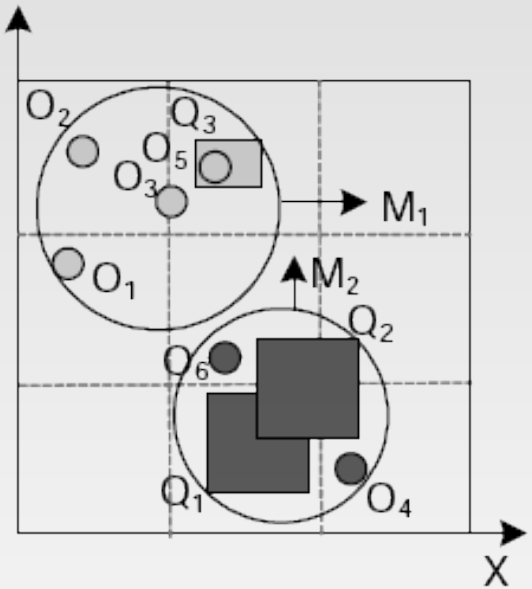
SCUBA state diagram:



SCUBA example (1/2)

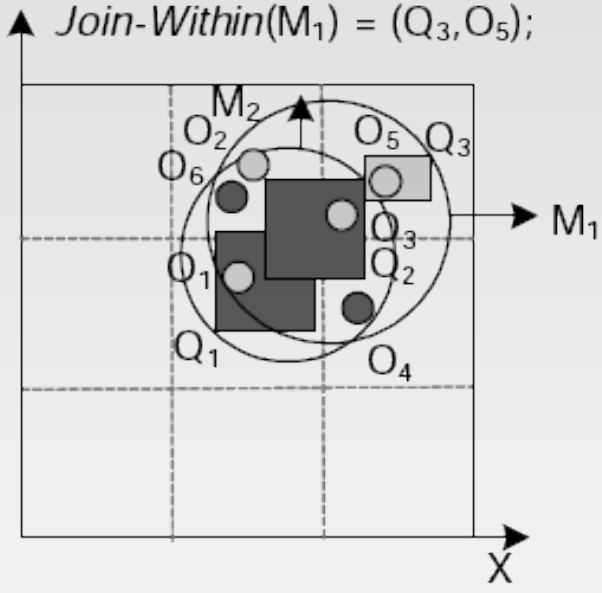


SCUBA example (2/2)

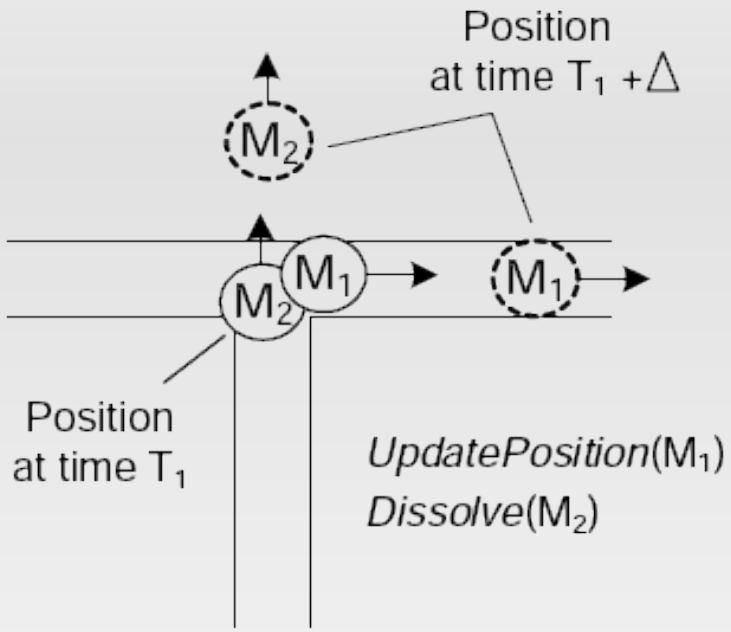


(d) Color-differentiating moving clusters M_1 and M_2

Results:
Join-Between(M_1,M_2) = *overlap*;
Join-Within(M_1 U M_2) = (Q_2,O_3);
Join-Within(M_1) = (Q_3,O_5);



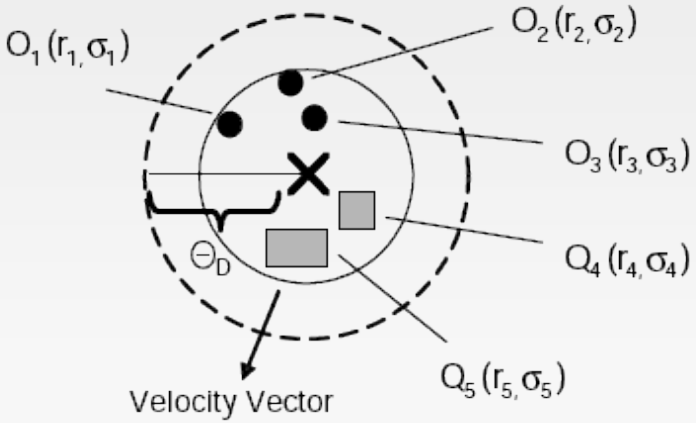
(e) Snapshot at time T_1: Joining moving clusters M_1 and M_2

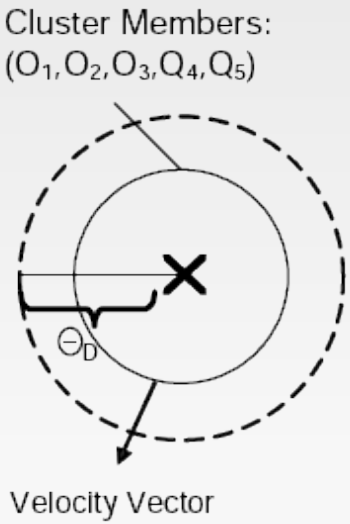


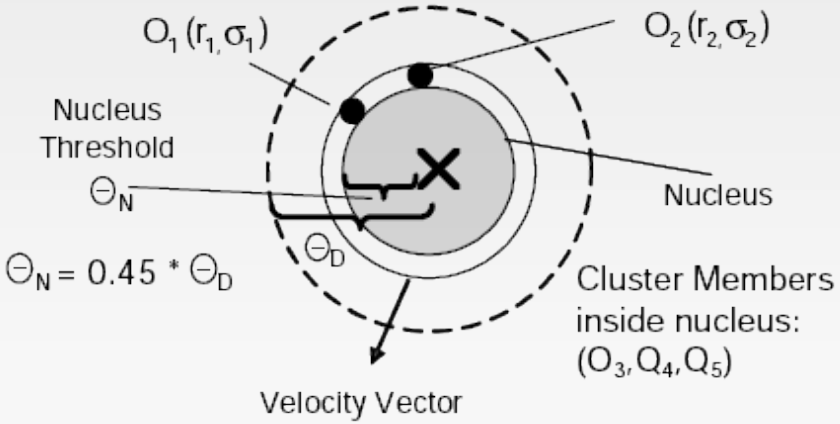
(f) Cluster post-join maintenance

Optimisation

- Moving Cluster based Load Shedding.
  - Can further decrease the amount of joins to perform
  - Random vs semantically less important.
  - Non, full or partial.

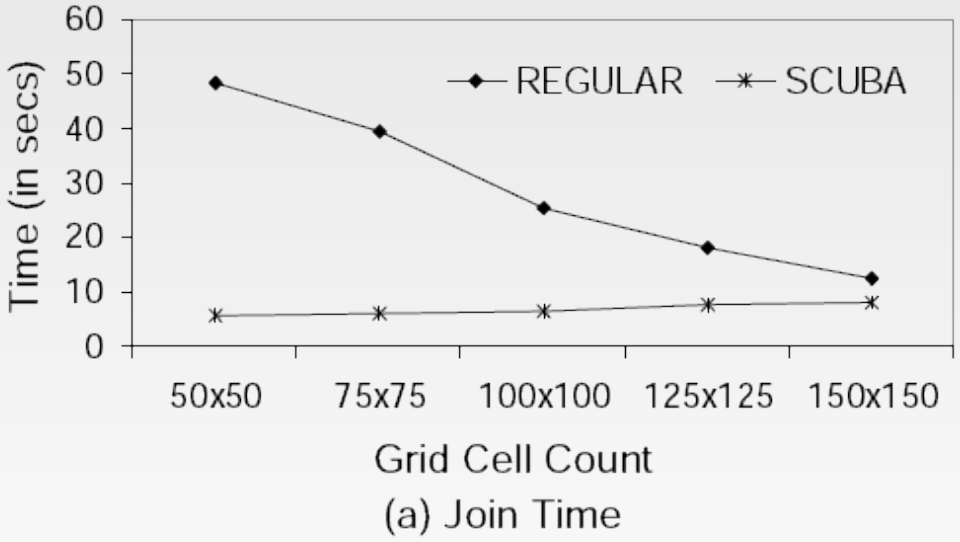




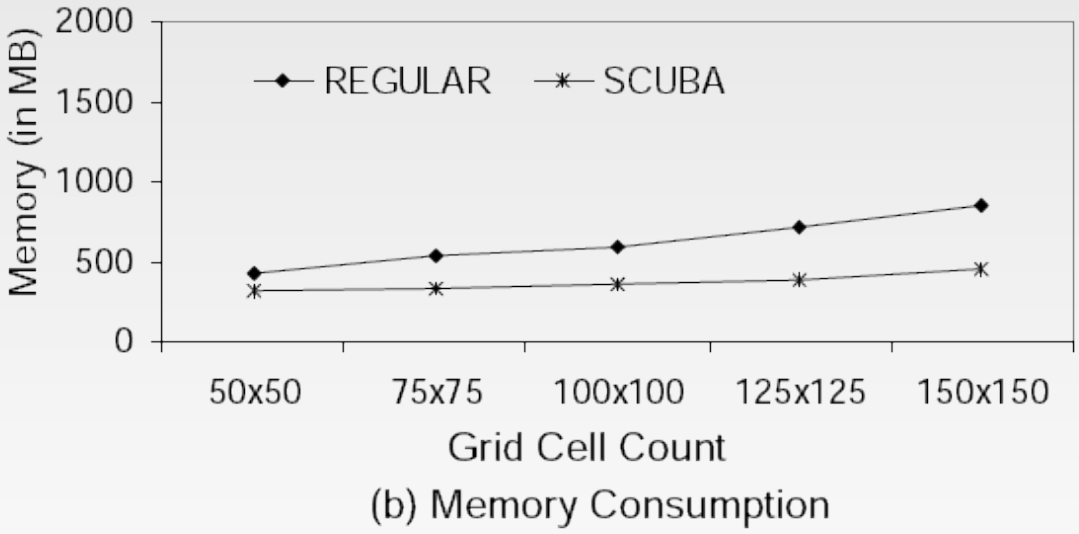


Results (1/3)

- The effect of grid cell size when comparing SCUBA to a regular grid based index.



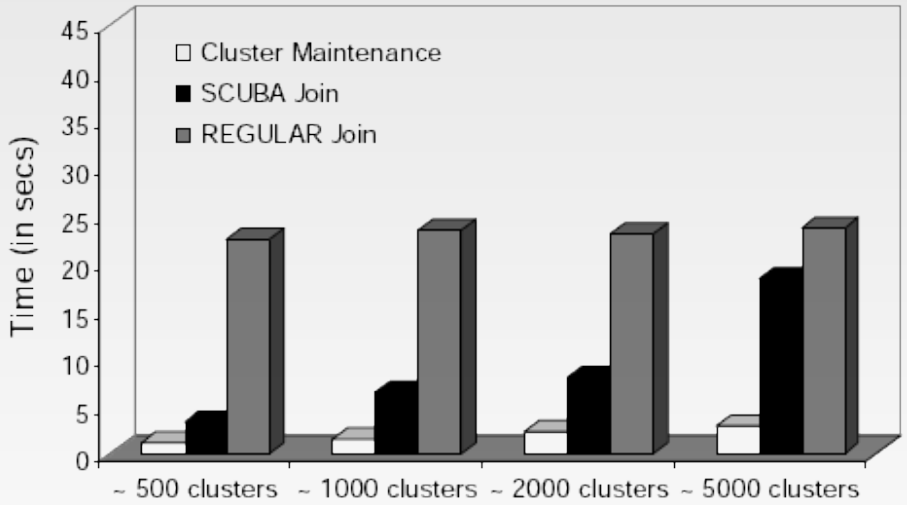
(a) Join Time



(b) Memory Consumption

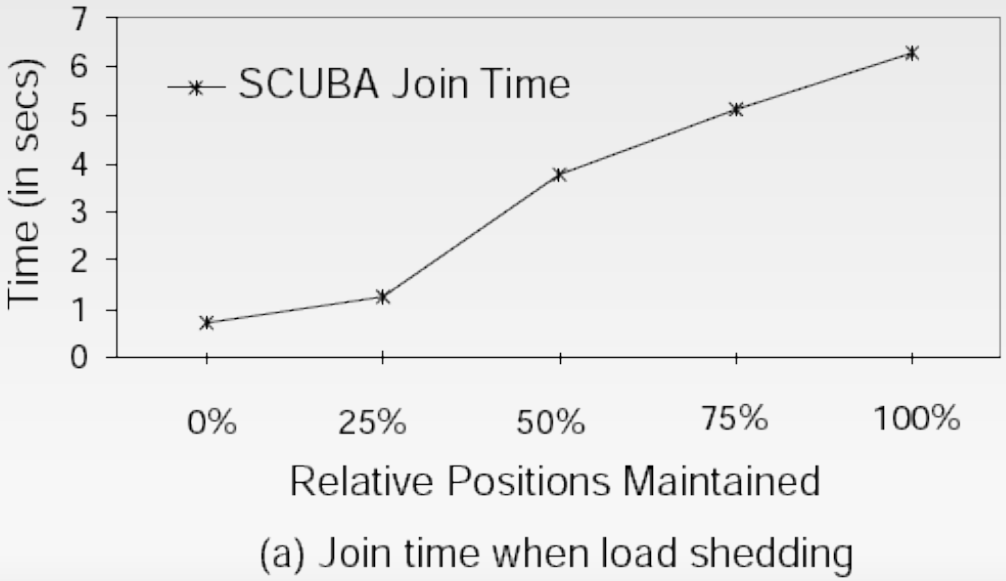
Results (2/3)

- The effect of maintenance in SCUBA.
  - Cluster maintenance is costly, but combined with the decrease in join time it still makes SCUBA faster than regular join in grid based index.

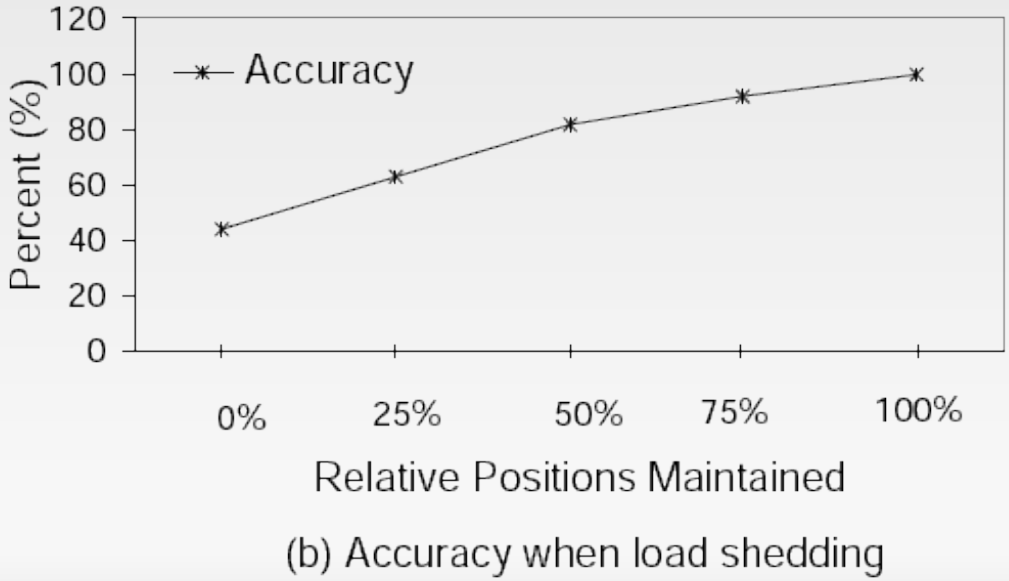


Results (3/3)

- Load shedding optimisation.
  - To decrease memory usage it may become necessary to activate Load Shedding. This test shows how different degrees of shedding affects the accuracy of the results.



(a) Join time when load shedding



(b) Accuracy when load shedding

Conclusion

- Facilitates efficient execution of queries on moving objects with similar spatio-temporal attributes.
- SCUBA can give significant optimisations to join time compared to the traditional grid based index approach.
- Load shedding can reduce execution time with minimal degradation in results.
- First to utilise motion clustering on spatio-temporal data streams.

Related work and contributions

- Dat5 project: Movement tracking of passengers in an airport by Bluetooth devices.
- Utilise some aspects of clustering of objects with similar attributes, to make real time passenger flow analysis.
- Can maybe be used in planned LBS in airport system.

Evaluation of paper

- Well written paper with a good “flow”.
- The theories are presented in a natural order and finalised with thorough well documented experimental results.
- Lacking more comparison to some of the other mentioned methods, e.g. Query Indexing.

That is all

Thank you for listening.

Any questions, comments?