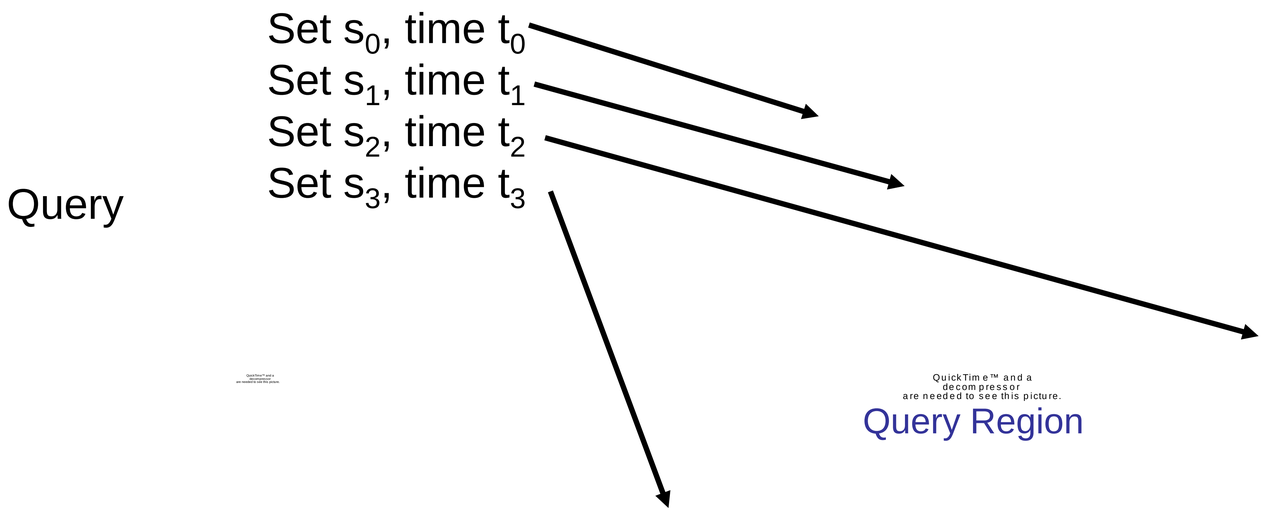
Persistent Spatial Data Structures

Stuart A. MacGillivray and Bradford G. Nickerson
Faculty of Computer Science, University of New Brunswick, Fredericton, New Brunswick, Canada

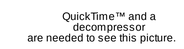
Problem Definition and Motivation

- Sets of spatial data points added at distinct time epochs.
- Range searches only return most recent data where sets overlap.
- Irregular update regions may require alpha-shapes for definition.
- Queries can optionally omit subsets from consideration.
- Motivation: Queries on time-stamped massive data surveys.
- Challenge: Massive data requires I/O model; complex queries.



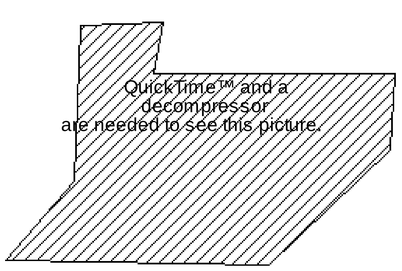
Persistent Data Structures

- Persistent data structures maintain versioning history and data, retaining search complexity.
- Alterations tracked between versions of structure.
- Partially (only current version editable) or fully (all versions editable) persistent.
- Planar point location in O(log n) query, O(n) space.
- Most persistent structures are non-spatial.



Testing Plans

- Data sets: Shallow Survey 2008 Common Dataset (CCOM/JHC), as well as a synthetic data set.
- SSCD contains >580 GB of survey data; sets of binary point data range from 2 to 28 GB.
- Synthetic data set for initial testing randomly generated.



Planned Approaches

- Can simply treat time as an additional dimension.
- Requires complex query for stated results.
- Preferable: Make a persistent spatial data structure.
- Time treated as discrete epochs, making this approach more applicable.
- Can a 2-d data structure be made persistent to support rectangle + time epoch queries in O(log N + K) time using O(N) space?
- Can updates be performed in O(M log N) time for M points added at epoch t_i?
- Can persistent spatial data structures be made I/O-efficient?

References:Peyman Afshani, Lars Arge, and Kasper Dalgaard Larsen. Orthogonal range reporting in three and higher dimensions. In FOCS, pages 149-158. IEEE Computer Society, 2009.
James R. Driscoll, Neil Sarnak, Daniel Dominic Sleator, and Robert E. Tarjan. Making data structures persistent. J. Comput. Syst. Sci., 38(1):86-124, 1989.
Neil Sarnat, Robert E. Tarjan. Planar Point Location Using Persistent Search Trees. CACM, 29(7):669-679, 1986.
Shallow Survey 2008. http://www.shallowsurvey2008.org/

Sponsored by:
Faculty of Computer Science UNB
NSERC CRSNG