Exclusion Persistence in Spatial Data

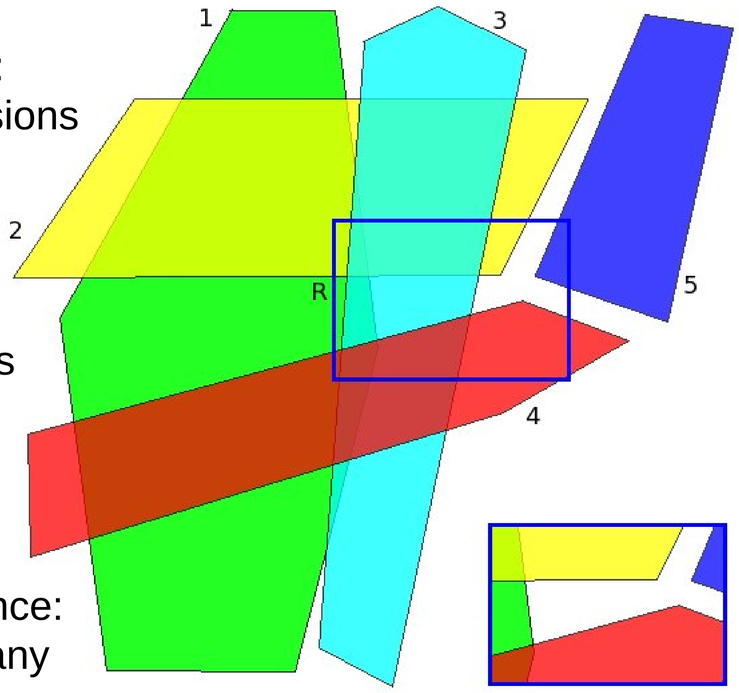
Stuart A. MacGillivray and Bradford G. Nickerson

Faculty of Computer Science, University of New Brunswick, Fredericton, New Brunswick, Canada

- Motivation: Efficient search of massive geographically referenced overlapping data sets

Persistent Data Structures

- Persistent data structures maintain version history.
- We examine a set of N points added over m updates, assuming m << N.
- Partial persistence: Queries on past versions possible, can only modify the most recent version.
- Full persistence: Edits to past versions create alternate branches, forming a tree of versions.
- Exclusion persistence: Queries can ignore any subset of past updates.

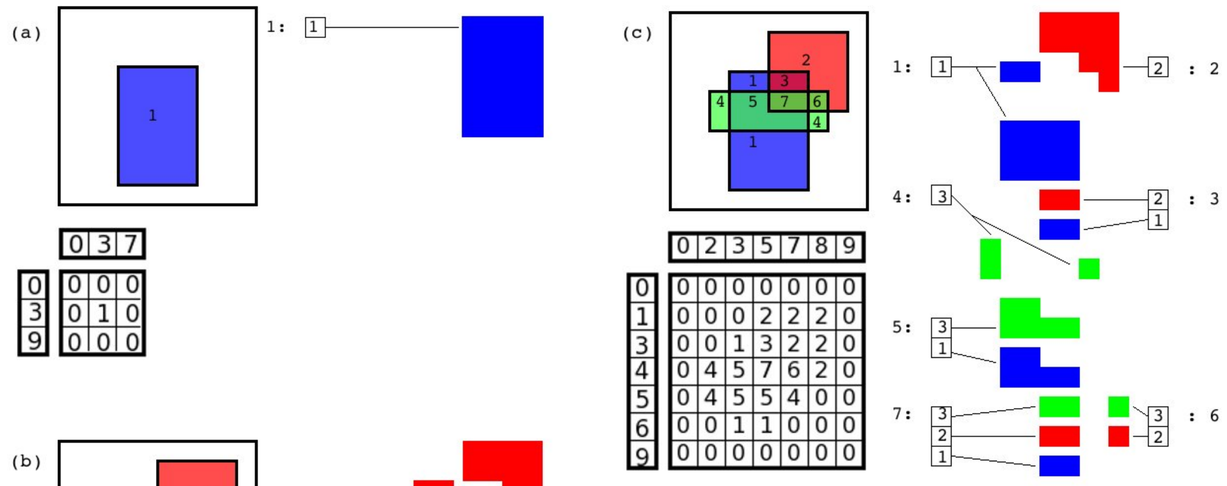


Q = (R, 5, {3}) returns

Current Results

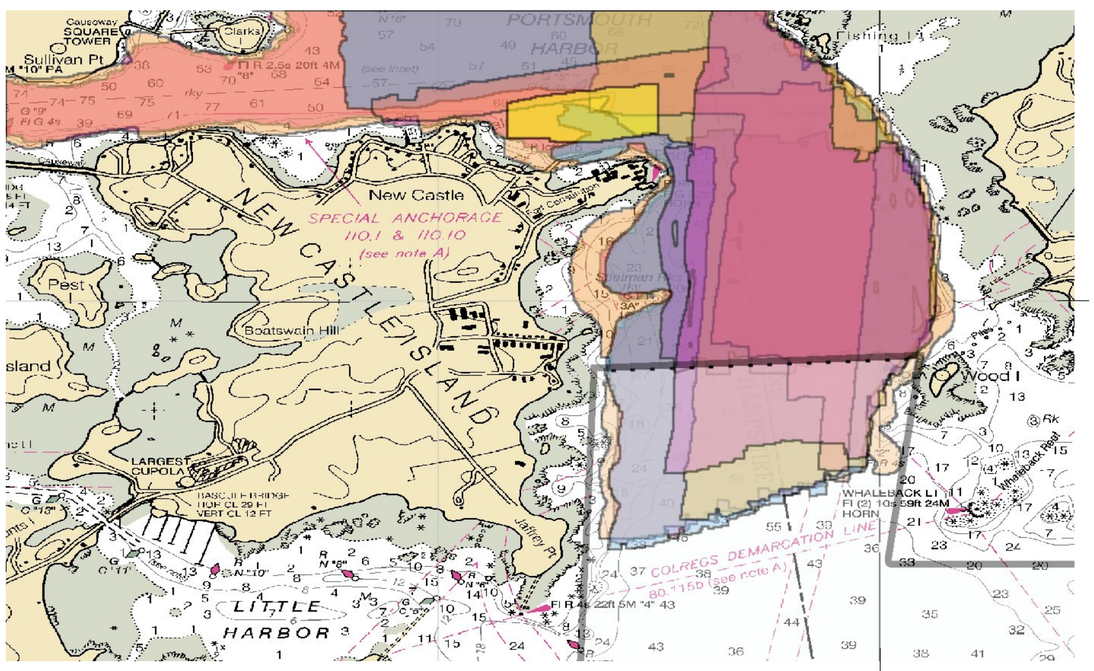
- An exclusion persistence range search is defined as $Q = (R, t_q, T_e)$ where t_q is the index of the query time and T_e is the set of time indices whose matching updates are ignored.
- Challenge: Minimizing storage space and query I/Os at once.
- Current result: Overlapping 2-d data sets searchable in $O((mN/B)^{1/2} + fm^2 + K/B)$ I/Os to return K points, assuming block size B [1]. Storage requires linear space, by using stack-based indexing of overlapping f-sided convex polygons.
- Indexing of rectangular regions is possible with $O(m^2)$ memory.

Data Structure for Rectangular Indexing



- Index array stores subregion extents and enumerates them.
- Update extents are rectangular, but subregions can be non-convex and even disconnected.
- Each subregion is associated with a stack of pointers to linear space I/O-efficient optimal data structures.
- Step (b) has 2 updates, 3 subregions, and 4 data structures; Step (c) has 3 updates, 7 subregions, and 12 data structures.
- Each data structure stores points added to the associated subregion at the associated time, even if disconnected.
- Range queries search at most one structure per subregion.

Test Data, Shallow Survey 2008, m = 8



From Shallow Survey 2008. http://www.shallowsurvey2008.org/

References: [1] Stuart A. MacGillivray, Bradford G. Nickerson: A Data Structure Supporting Exclusion Persistence Range Search. Proceedings of the 24th Canadian Conference on Computational Geometry, CCCG 2012, Charlottetown, Prince Edward Island, Canada, August 8-10, 2012, pages 29-34

Sponsored by: UNB Faculty of COMPUTER SCIENCE; NSERC CRSNG