TOUCH: In-Memory Spatial Join by Hierarchical Data-Oriented Partitioning

A. Logins

Moscow Institute of Physics and Technology
Skolkovo Institute of Science and Technology

Course: Machine Learning and Data Analysis
(Strijov's practice)/Group 174, 2014 Fall

Goal of research

Motivation

Develop In-Memory Spatial Join algorithm with Iterative Hierarchical Data-Oriented Partitioning with balanced workload

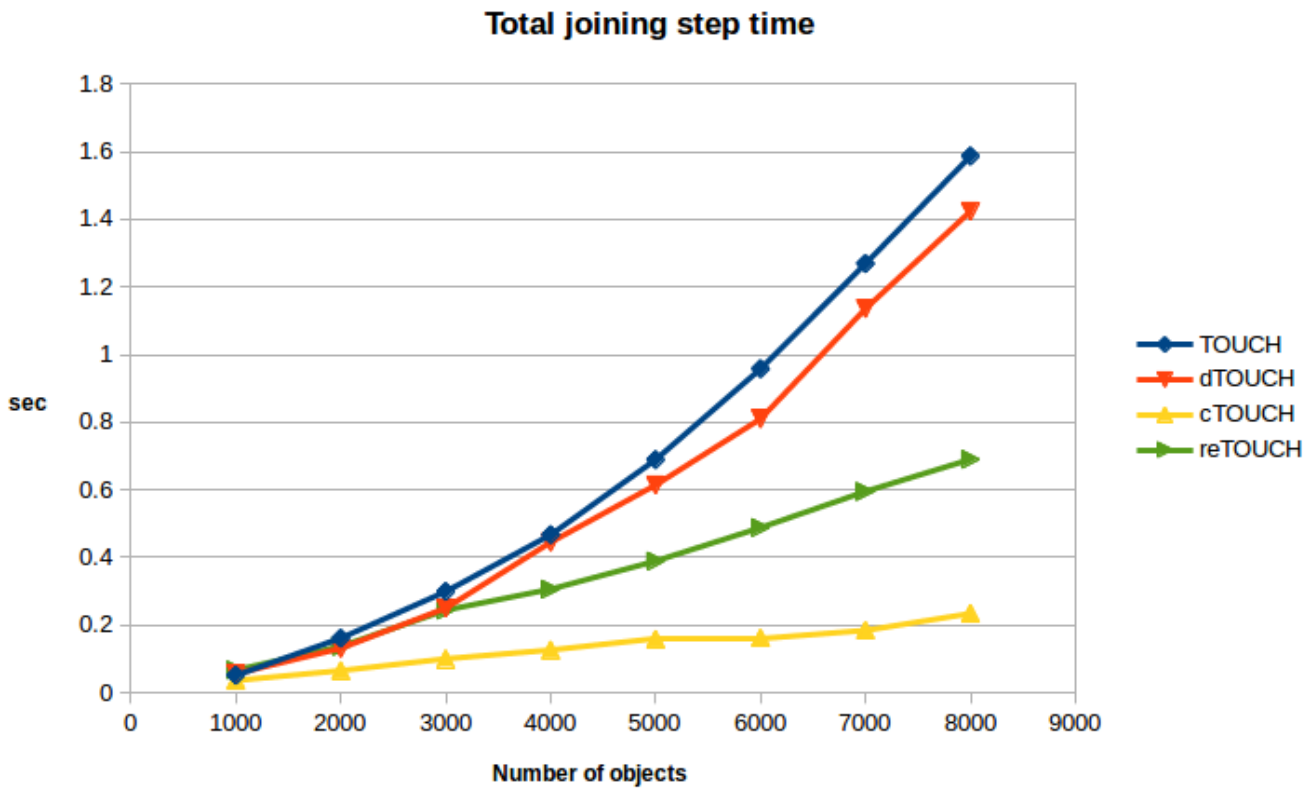
Problem statement

Given the parameter ϵ two datasets of spatial objects A and B find all $a \in A$ and $b \in B$ such that minimum distance between them is less than ϵ.

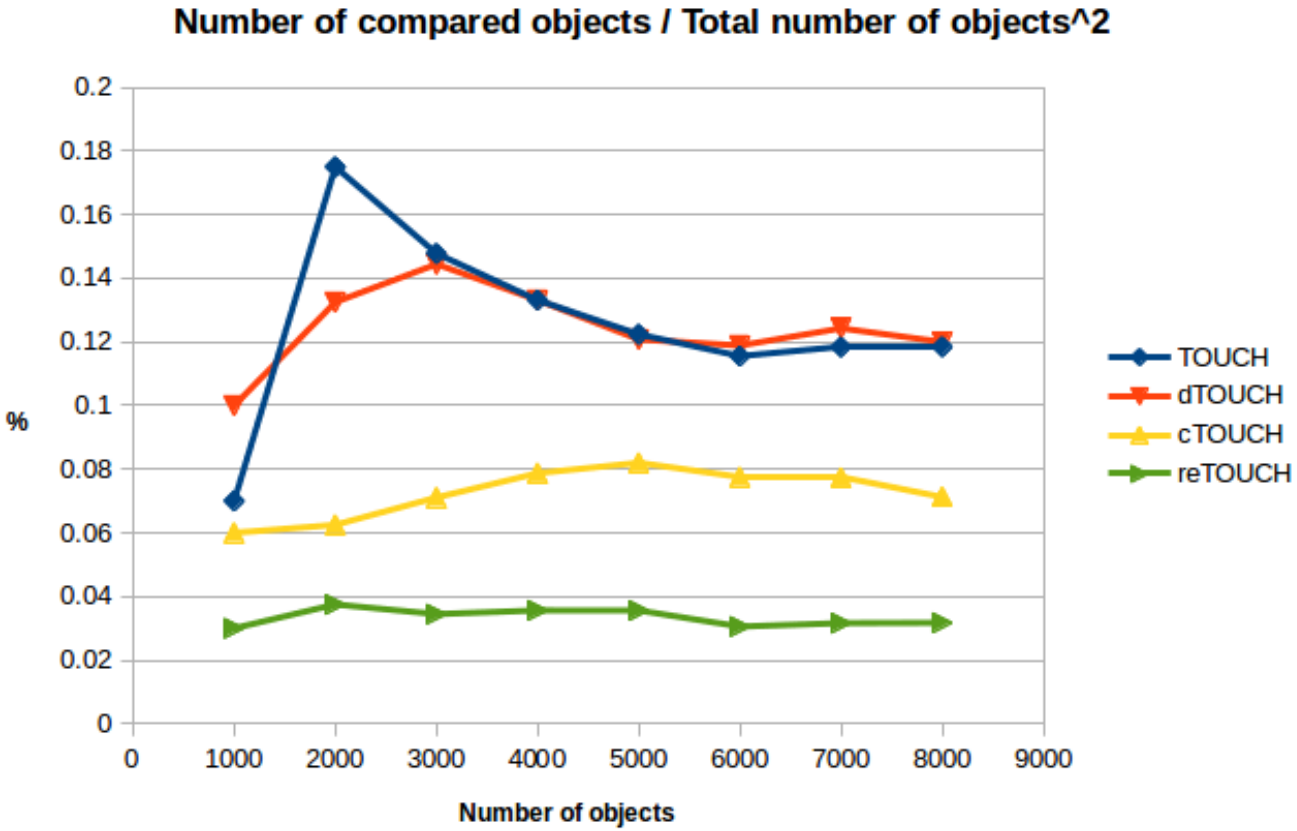
Solution

- Build Hilbert curve through all MBRs of spatial objects (create index)
- Build R-tree through indexed MBRs, maintaining MBRs of two types (according to types of objects) for each node
- For each leaf node take objects and assign to the nodes of the tree, dynamically updating MBRs and deleting them from the leaf nodes
- For each node join assigned objects with object assigned below

Computational experiment



Computational experiment



Conclusion

Two of three new approaches give considerable improvement in performance and number of comparisons.