A Disk-Based Index for Trajectories with an In-Memory Compressed Cache

Daniela Campos*, Adrián Gómez-Brandón†, and Gonzalo Navarro*

*IMFD, Dept. of Computer Science
University of Chile, Santiago, Chile
{dcampos,gnavarro}@dcc.uchile.cl

†CITIC, Database Lab, Fac. Informática
Universidade da Coruña, A Coruña, Spain
adrian.gbrandon@udc.es

The increasing number of personal devices that contain GPS generates a large amount of information about free-space trajectories, which is challenging to store and analyze. Classical spatio-temporal indexes, like the MVR-tree [1], store those trajectories on disk while supporting efficient searches. Various recent approaches, like GraCT [2], use compression to maintain the trajectory data in memory, which speeds up queries significantly. Compression, however, typically loses precision (e.g., GraCT discretizes the positions in the trajectories), which might not be acceptable.

In this work we explore the use of GraCT as an *in-memory lower-precision cached index* for a full-precision index on disk. The conservative answers (i.e., with false positives but no false negatives) returned by GraCT are thus checked against the full-precision data stored on disk, which requires a much simpler disk-based index: A B-tree storing the timestamps and positions of each object, with internal nodes storing MBRs of the descendant positions.

Our structure is shown to be time-competitive with the MVR-tree in time-slice (return all objects inside a given spatial window at a given time instant) and kNN queries (return the k objects closest to a given point at a given time instant), and sharply outperforms it in time-interval queries (return all objects inside a given spatial window at some moment in a given time interval). In addition, our B-trees can easily retrieve full or partial trajectories of given objects, which is not simple with the MVR-tree. Further, because its on-disk index is much simpler, our structure uses 24 times less disk space than the MVR-tree.

References

[1] Yufei Tao and Dimitris Papadias, "MV3R-tree: A spatio-temporal access method for timestamp and interval queries," in *VLDB*, 2001, pp. 431–440.

[2] Nieves R. Brisaboa, Adrián Gómez-Brandón, Gonzalo Navarro, and José R. Paramá, "GraCT: A grammar-based compressed index for trajectory data," *Information Sciences*, vol. 483, pp. 106 – 135, 2019.

Supported by Xunta de Galicia/FEDER-UE under Grants [IN848D-2017-2350417; IN852A 2018/14; ED431C 2017/58]; Xunta de Galicia and European Union (European Regional Development Fund- Galicia 2014-2020 Program) with the support of CITIC research center under Grant [ED431G 2019/01]; Ministerio de Ciencia, Innovación y Universidades under Grants [TIN2016-78011-C4-1-R]; Also supported by ANID - Millennium Science Initiative Program under Grant [ICN17_002].