Integration of Mobility Data with Weather Information

Nikolaos Koutroumanis, Georgios M. Santipantakis, Apostolos Glenis, Christos Doulkeridis, George A. Vouros
koutroumanis@unipi.gr, gsant@unipi.gr, apostglen46@gmail.com, cdoulk@unipi.gr, georgev@unipi.gr

Department of Digital Systems, University of Piraeus, Greece

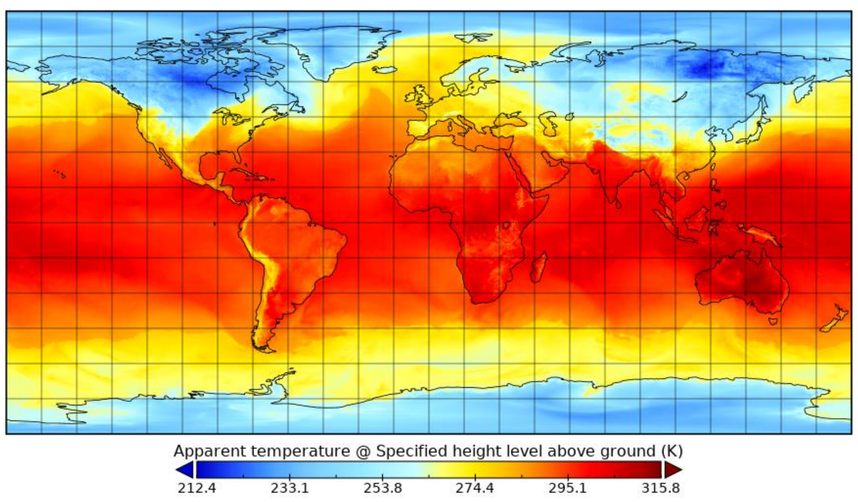
1. Weather Data Integration Task

Objective:

➢Enrich mobility data with weather information by associating spatio-temporal positions of moving objects with external data sources.

Data sources with weather information

➢We use the GRIB files based on the Global Forecast system (GFS) offered by National Oceanic and Atmospheric Administration (NOAA), which is a type of Numerical Weather Prediction data model and has the globe partitioned per 0.5 degrees.



Visualization of a weather attribute with Panoply

Contributions:

➢We propose a generic system for integrating mobility data represented by spatio-temporal positions.

➢We show how to extend the basic mechanism in order to perform weather integration for more complex geometries (3D large sectors).

➢We demonstrate the efficiency of our system on real mobility datasets from different domains (urban, maritime and air-traffic).

2. System Architecture

Components:

Spatio-Temporal Parser

➢Parses sequentially the records of input data set of mobility data

➢For each record, a set of basic cleaning operations are performed (null of empty values and logitude – latitude validity)

Weather Data Obtainer

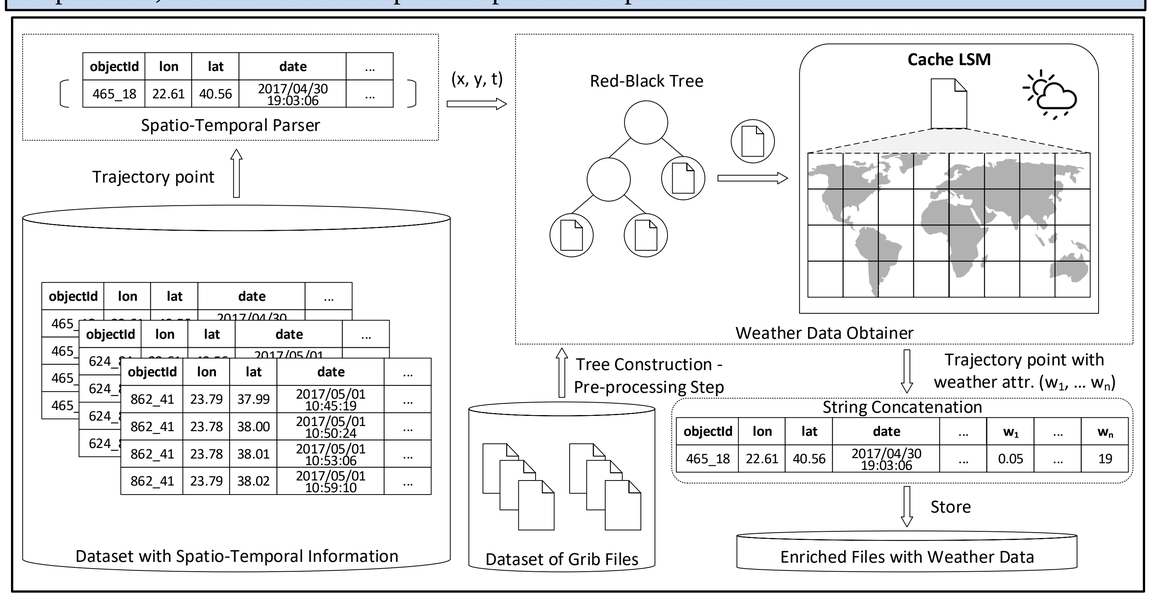
➢Takes as input Spatio-temporal information (*x*, *y*, *t*)

➢Determines the right file that should be accessed in order to get the values of the weather attributes by using Red-Black Tree.

➢Exports the record with the added weather attributes

Caching:

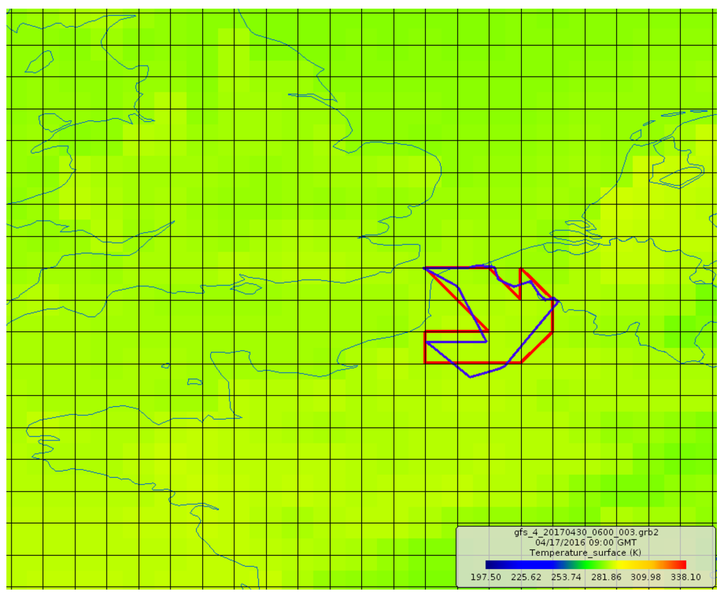
We introduce a simple caching mechanism, which practically maintains in memory references to open files, in order to avoid repeated open/close operations.



3. Extensions

1) Enabling Complex Geometries with Weather Data

- In many cases we may need to associate the trajectory of a moving object (i.e. a LineString geometry) with weather conditions, or a region or a cluster of regions on the surface of the earth (i.e. polygon or multi-polygon geometries).
- Given a geometry g given a geometry and a time interval [ts ,te] the average of all the values retrieved for all the points of the geometry are returned.



Example of airblock (in blue), its simplified geometry (in red), and temperature surface shown in the coloured map.

2) Providing Linked Data in RDF

- Since RDF is the W3C standard to be used for Linked Open Data, connecting as a consumer to RDF triples, it can exploit any positioning data available on the web, to return it enriched with weather data.

4. Experimental Evaluation

Experimental Setup:

➢Centralized computer with i7-4790 processor, 16GB RAM, 1TB HDD and Ubuntu OS

➢Dataset: Real mobility data containing real trajectories of vehicles in the region of Greece

➢13 weather attributes (rain-related) have been chosen to be enriched in the dataset

Evaluation on large data set

Weather Integration	Execution Time	Memory Consumption	Throughtput
With Indexing	12 sec	229 MB	3,570 rows/sec
Without Indexing	1,660 sec	106 MB	26 rows/sec
Pre-processing	1 sec	57 MB	N/A

Caching Mechanism Evaluation (case of randomly shuffled records)

Throughput for increased cache size

Software available for download: https://github.com/nkoutroumanis

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687591, No 780754, No 779747 and No 777695.