FREQUENT ROUTE BASED CONTINIOUS MOVING OBJECT LOCATION- AND DENSITY PREDICTION ON ROAD NETWORKS

Győző Gidófalvi, Manohar Kaul, Christian Borgelt, and Torben Bach Pedersen

ROYAL INSTITUTE OF TECHNOLOGY
European Centre for Soft Computing
UPPSALA UNIVERSITET
AALBORG UNIVERSITY DENMARK

Problem Statement

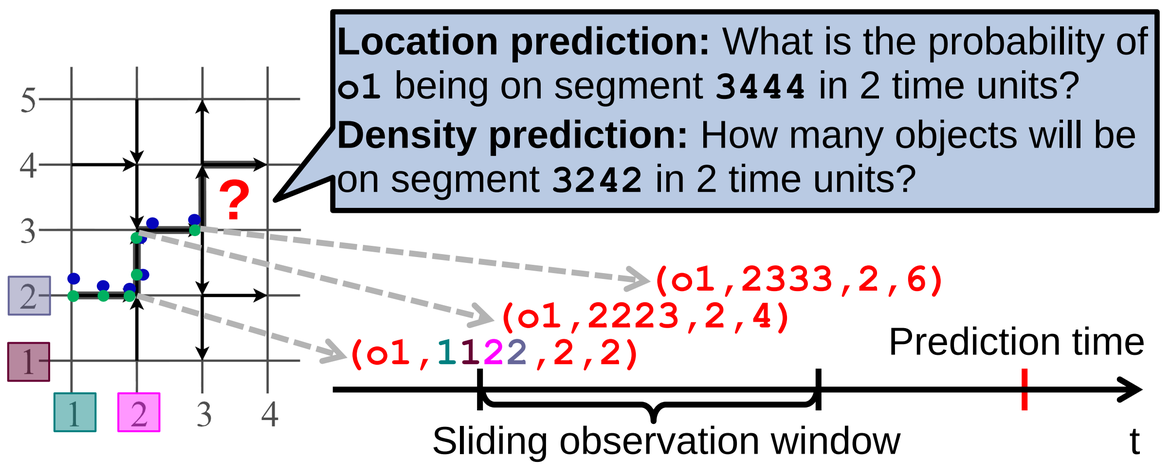
Given the current and historical movements of vehicles, predict their near-future location on the road network to:

1. Estimate near-future traffic conditions: (density/speed/flow)
2. Provide actionable travel information based on future estimates

- Inform the relevant vehicles in case of an (actual/predicted) event
- Suggest how and which vehicles to re-route in case of an event

Preliminaries

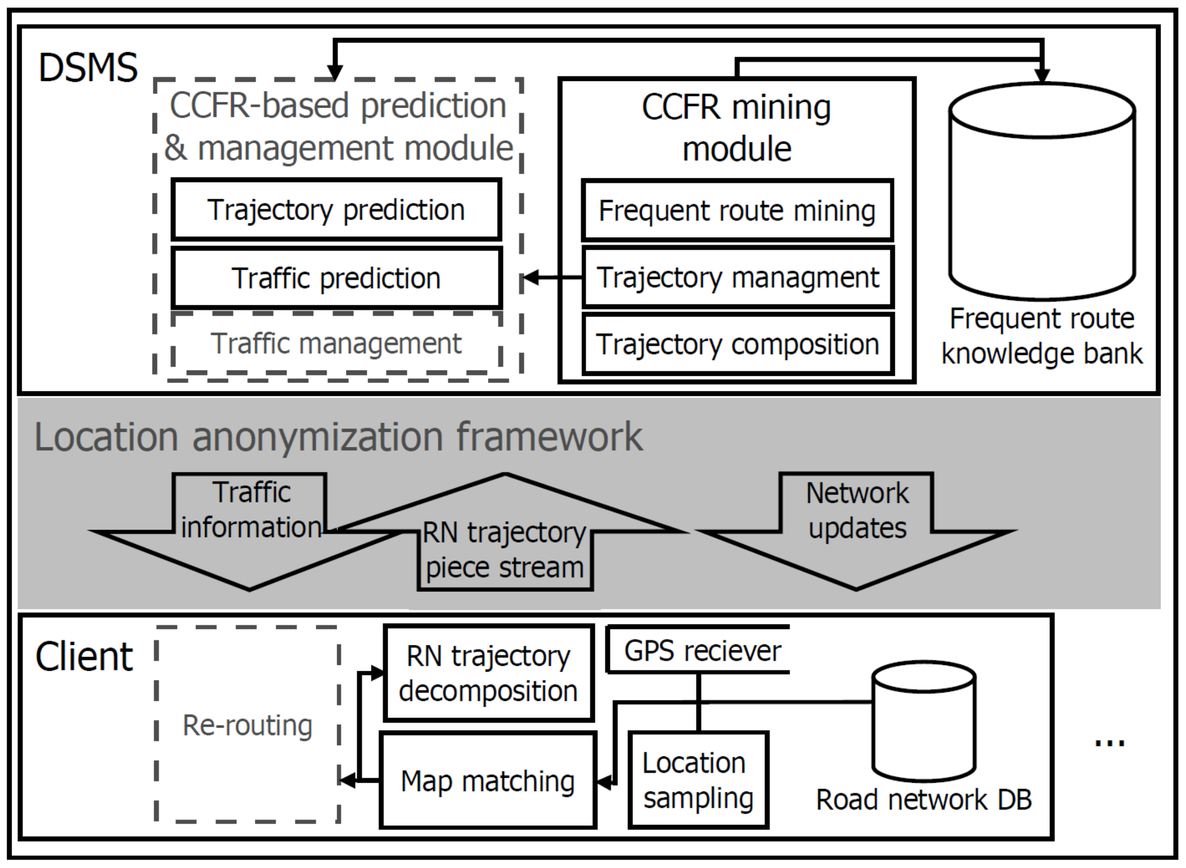
Movement representation: Map matching clients align noisy GPS measurements to precise RN locations. When an RN segment is fully traversed a *trajectory piece* `(oid,segment,traversal time,[arrival_time])` is sent to the sever, i.e., the RN trajectory is a sequence of trajectory pieces. *Stops* subdivide RN trajectories into a discontinuous sequence of *trip trajectories*. Object movement is observed as a *stream of evolving trip trajectories*.



Knowledge representation: **Closed Contiguous Frequent Routes (CCFR)**:

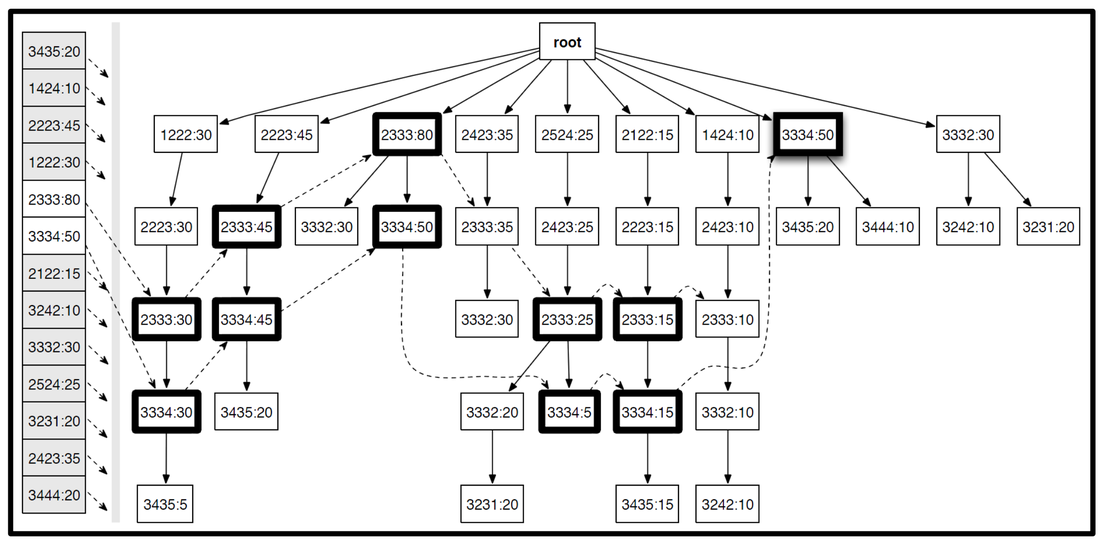
- **Closed** → lossless compression of knowledge
- **Contiguous** = ”no gaps” → effective prediction

Architecture



CCFR-Based Prediction

1. **Mine and store CCFRs continuously.**
   Depth-first search
   Closedness: direct check of pattern extension
   Also calculate *turn statistics*
2. **Retrieve and insert previously mined, relevant CCFRs in an FP-tree *τ*.**



3. **Given an object’s current trajectory *qv*:**
   I. Find the branches that ”best” match *qv* in *τ*.
   (NOTE: Turn statistics guarantee a match)
   II. Calculate the probability of each possible next segment from the *closed* pattern supports.
   III. Distribute the probability mass of the object proportional to the segment probabilities, extend *qv*, and recurs until the time horizon is reached.
   IV. Aggregate the predicted object locations at the time horizon → predicted network density.

Empirical Evaluation

Hardware: Macbook Pro, Intel Core 2 Duo 2.4 GHz CPU, 4MB L2 Cache, 800 MHz Bus speed, and 4GB memory

Data set:

- Real-world trajectories of 1500 taxis and 400 trucks in Stockholm
- Road network: 6000 directed, 55 meter long segments with a connectivity of degree of 2.3
- 17000 trip trajectories during the course of a day

Results

Throughput and scalability: Even for large mining windows (24 hours ≈ 17K trajectories) and very low support values (0.1%) the execution time is within real-time processing limits (<1 minute). Scales nearly linearly with the number of trajectories.

Prediction accuracy: Outperforms the “turn statistics only" approach by 10-30%. The additional prediction utility of the proposed approach becomes increasingly pronounced as the time horizon is increased.

Győző Gidófalvi: KTH Royal Institute of Technology – Geodesy and Geoinformatics – gyozo.gidofalvi@abe.kth.se
Manohar Kaul: Uppsala University – Department of Information Technology – manukaul@acm.org
Christian Borgelt: European Centre for Soft Computing – christian@borgelt.net
Torben Bach Pedersen: Aalborg University – Department of Computer Science – tbp@cs.aau.dk