Efficient One-click Browsing of Large Trajectory Sets

(Demo paper)

Benjamin Krogh, Edwin Lewis-Kelham, Ove Andersen, Kristian Torp

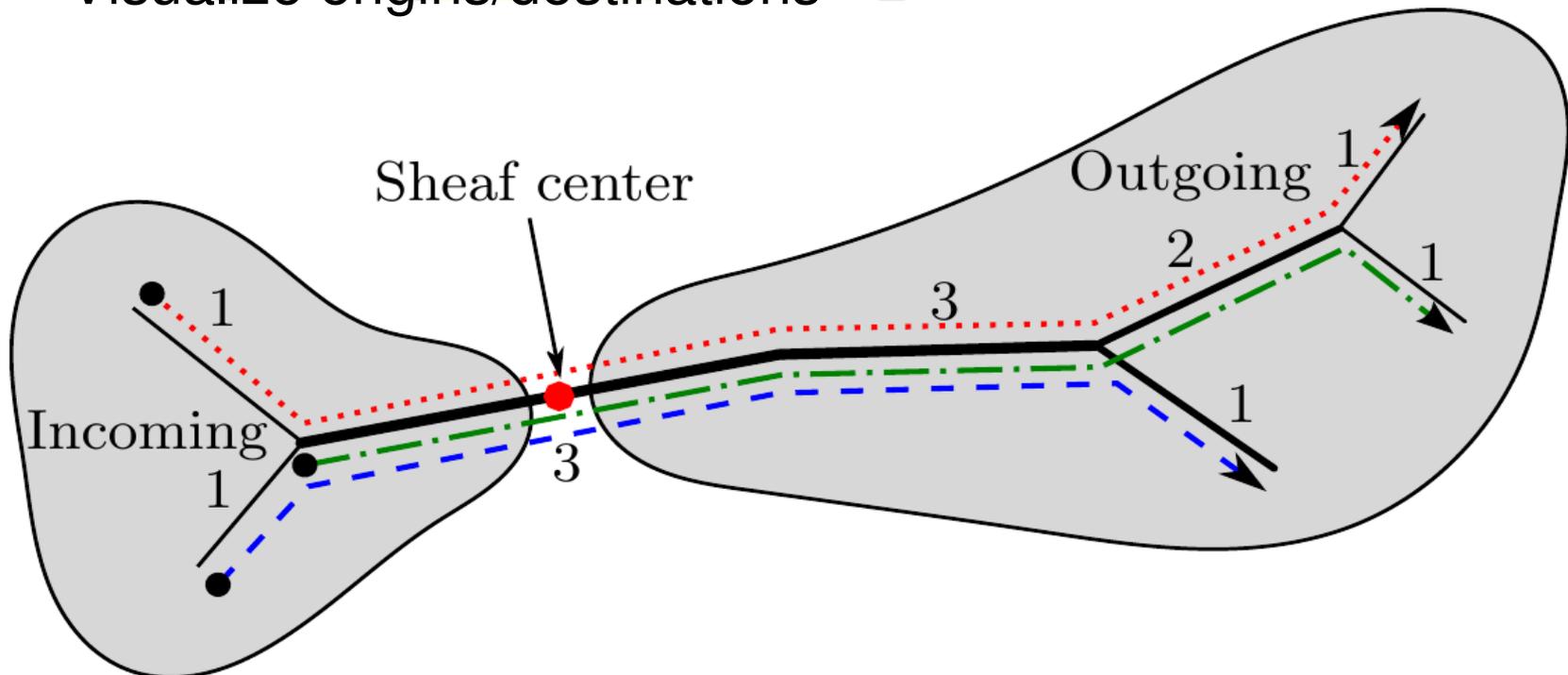**Department of Computer Science**

Aalborg University, Denmark

{bkrogh, edwin, xcalibur, torp}@cs.aau.dk

Sheaves – Basic Idea

- Place a *sheaf center* in the road network

- Retrieve all trajectories passing this center

- Visualize origins/destinations

Sheaves – Problem

- Retrieving all trajectories is expensive!

  – Trajectories may be very heterogenous

  – Up to 40K trajectories needs to be retrieved and processed for one sheaf

- Results in a random access pattern, difficult to index on disk

- Solution: Store everything in main memory!

Sheaves – Solution

- Memory is a scarce and expensive resource

- Trajectory data sets may be very large (ours has 2 million trajectories!)

- Compress data to reduce memory load

  – However, the individual uncompressed trajectory is very small (552 bytes, on average)

  – Difficult to provide random access while preserving compression ratio

Trajectory Compression - Spatial

- Observation: Trajectories follow the shortest path through the road network.

- Idea: Store only the endpoints of shortest paths.

- Path = [e1, e2, e3, e4, e5, e6, e7, e8, e9]

- Compressed path = [e1, e9]

    – Assuming [e2, e3, e4, e5, e6, e7, e8] is a shortest path

- Very good compression factor for real data (> 16)
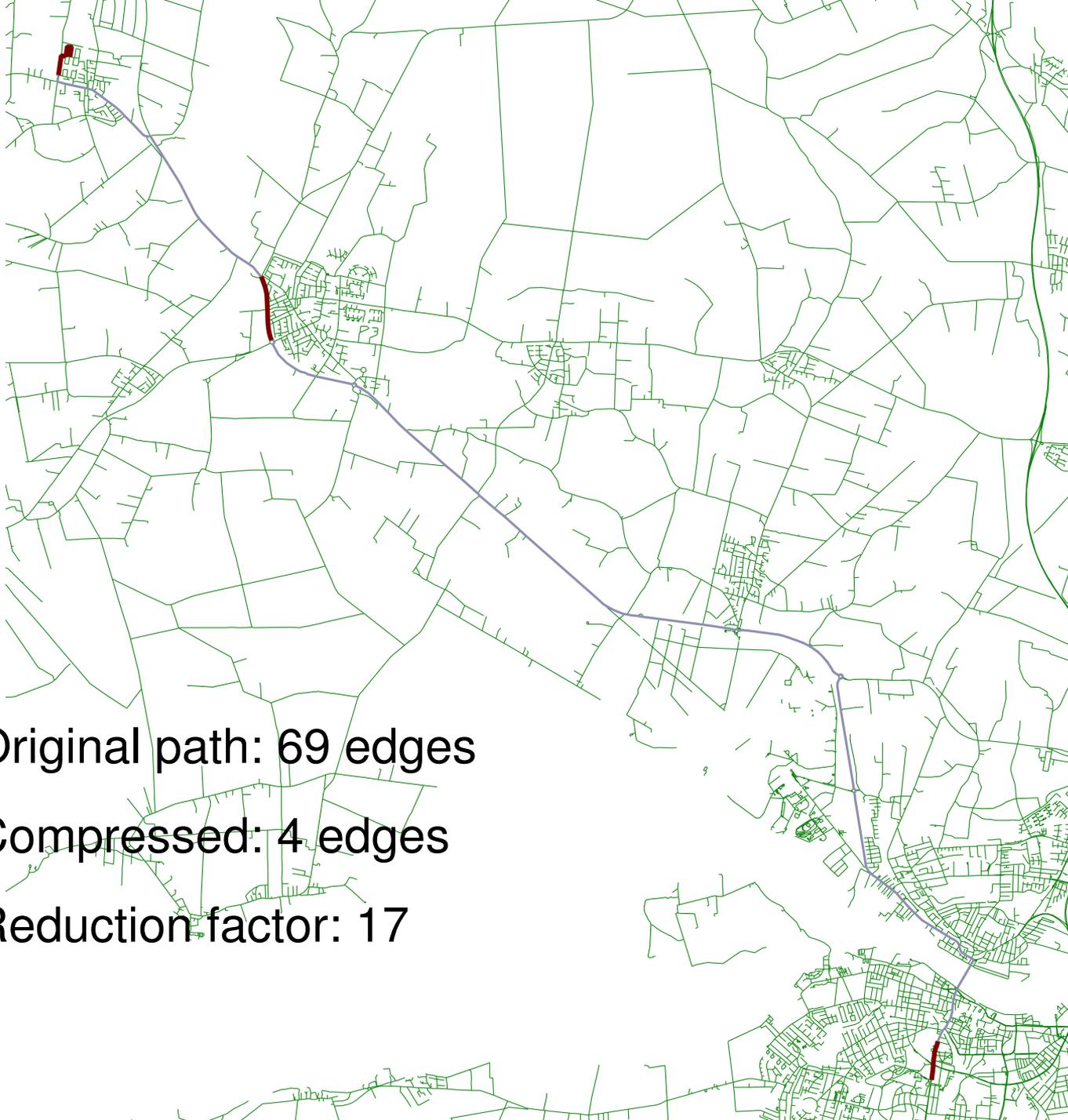
- Original path: 69 edges
- Compressed: 4 edges
- Reduction factor: 17

Trajectory Compression - Spatial

- Decoding of compressed trajectories requires computing many shortest paths (SPs).

  – Up to 120K shortest paths

  – Classical SP algorithms are too slow for this purpose

- The Hub Labeling (HL) SP algorithm requires < 1 microsecond per SP

  – Decoding of 40K trajectories using HL on quad-core i7: < 100 ms.

Trajectory Compression - Temporal

- Path = [e1, e2, e3, e4, e5, e6, e7, e8, e9]

- Timestamps = [ts0, ts1, ts2, ts3, ts4, ts5, ts6, ts7, ts8, ts9]

- Temporal information:

  – Start and end time for each edge touched

  – Essentially, a list of monotonically increasing integers

- Delta compression techniques for integers have very fast decoding speed (billions of integers/second)

Trajectory Compression

- Random access to compressed trajectories!

- Decoding is very fast!

- More than an order of magnitude reduction

Data type	Original	Compressed	Ratio
Spatial	497 MB	30 MB	16.6
Temporal	497 MB	53.7 MB	9.3

Summary

- Very high compression rate for network constrained trajectory data

- Interactive browsing of very large trajectory sets

- Use cases

  - Isochrones

  - Reachability analysis

  - Location based analysis

  - Intersection analysis

  - Analysis of changes in traffic over time

Acknowledgements