# On k-Path Covers and their Applications

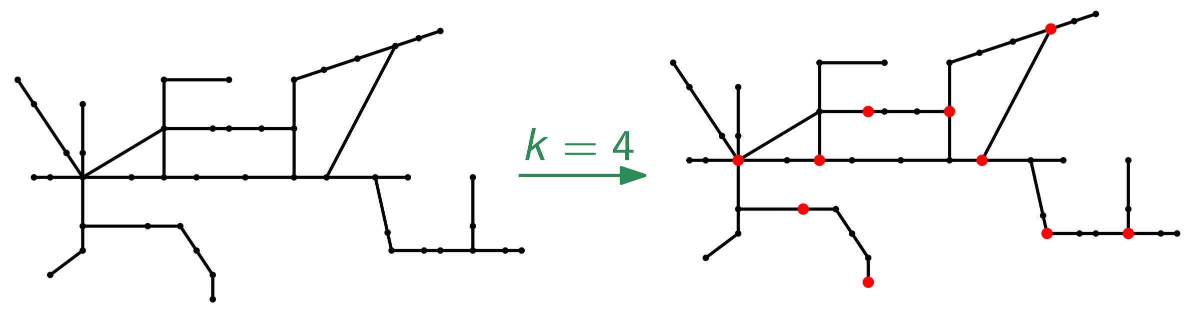
University of Stuttgart
Germany

Stefan Funke **and** André Nusser **and** Sabine Storandt
Universities Stuttgart/Freiburg, Germany

UNI FREIBURG

## PROBLEM DEFINITION

**Minimum k-(All-)Path Cover (k-APC)** Given a (di)graph $G(V,E)$ and $k \in \mathbb{N}$, select a minimum subset of vertices $C \subseteq V$ such that for every simple path $\pi = v_1, \cdots, v_k$ in $G$ we have $C \cap \pi \neq \emptyset$.




SPECIAL CASE

**Minimum k-Shortest-Path Cover (k-SPC)** Given a weighted (di)graph $G(V,E,c)$ with $c : E \to \mathbb{R}^+$, and $k \in \mathbb{N}$, select a minimum subset of vertices $C \subseteq V$ such that for every shortest (according to $c$) path $\pi = v_1, \cdots, v_k$ in $G$ we have $C \cap \pi \neq \emptyset$.

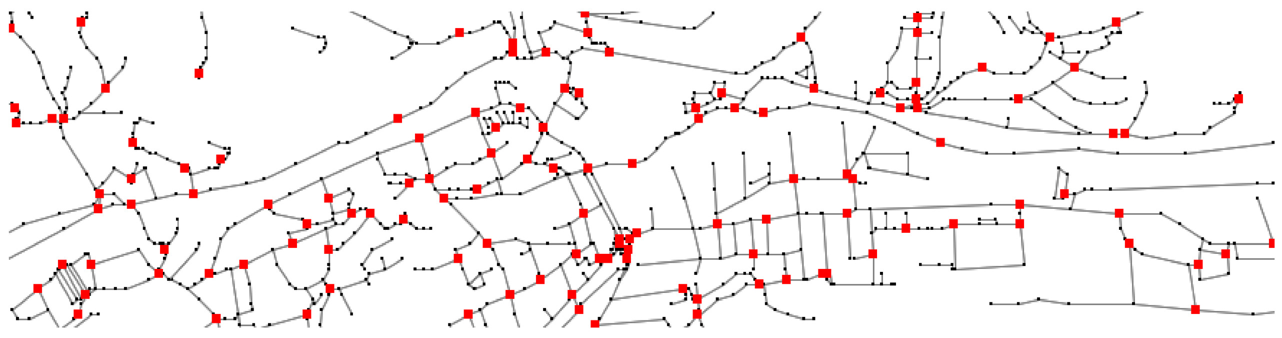
## APPLICATIONS

Handle large **Spatial Network Databases (SNDBs)** with billions of geographic entities.

### FACILITY LOCATION PROBLEMS

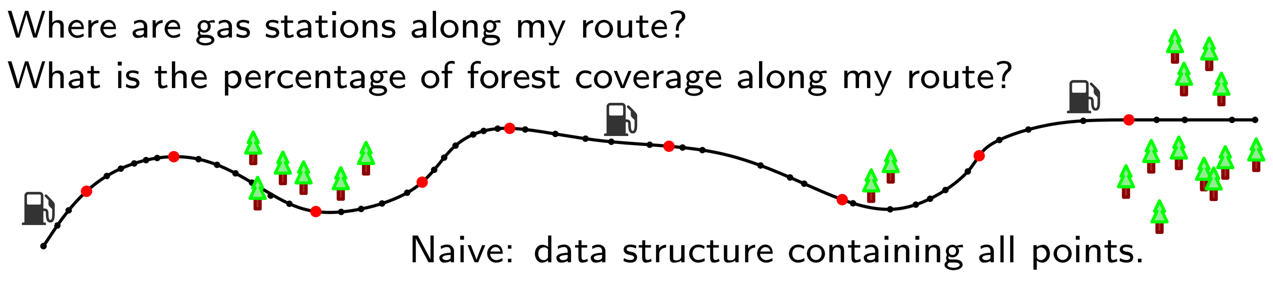
Place gas stations, signs, etc. such that every simple $k$-path is covered.

→ minimum k-APC



### DATA AGGREGATION & POI RETRIEVAL

Where are gas stations along my route?
What is the percentage of forest coverage along my route?

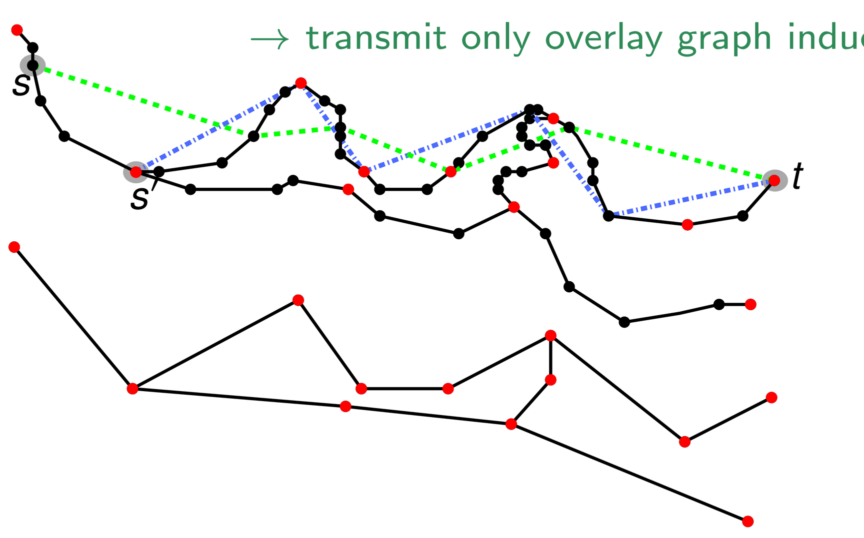


Naive: data structure containing all points.
→ data structure only on k-APC more efficient (space and time!)

### MAP SIMPLIFICATION

Transmit graphs and routes to a client for visualization.
Subsample to reduce bandwidth, but aim for consistent subsampling.

→ transmit only overlay graph induced by k-APC

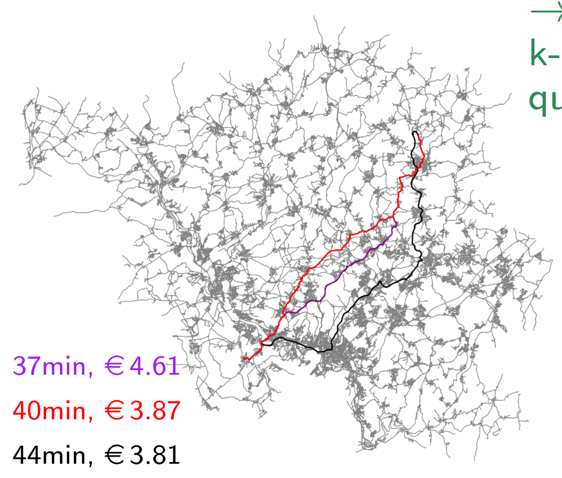


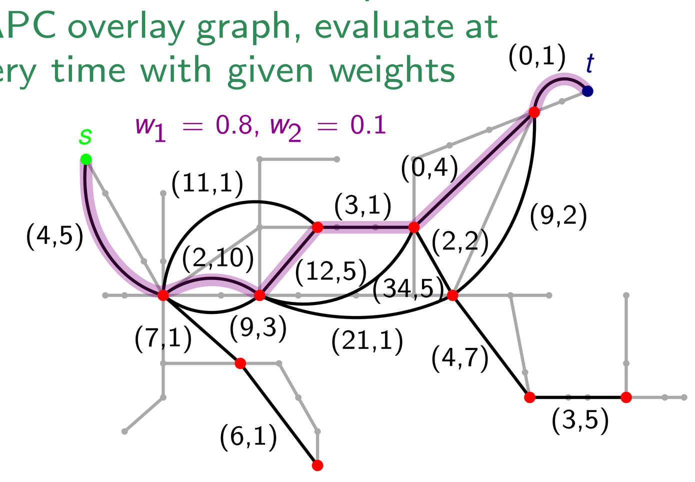

### PERSONALIZED ROUTE PLANNING

Metrics like travel time, gas price, scenic attractiveness, jam likeliness, etc. can be regarded to determine the optimal route between A and B.

Every user might have an indivdual set of weights (one for each metric) defining the importance of this metric for him. ⇒ revealed at query time

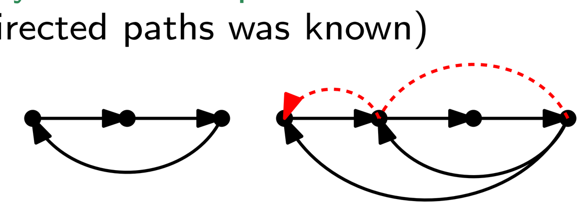
→ construct metric independent k-APC overlay graph, evaluate at query time with given weights




## THEORETICAL RESULTS

- k-APC and k-SPC are APX-hard (subsuming NP-hardness)
  ⇒ if the Unique Game Conjecture holds, k-APC and k-SPC cannot be approximated better than $2-\epsilon$
- k-approximation possible via pricing method (based on ILP formulation)
- $\log(OPT)$-approximation for k-SPC via VC-dimension analysis
  ⇒ New result: VC-dimension d of a system of unique directed shortest paths is 3 (before only $d=2$ for undirected paths was known)



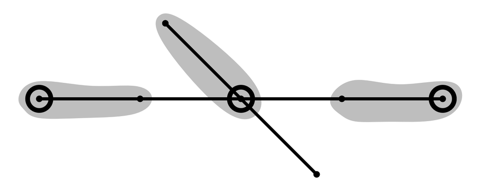
## PRACTICAL CONSTRUCTION

**Pruning Algorithm**

- start with $C = V$
- consider nodes one by one in some predefined order
- decide for each node whether it has to stay in $C$ to maintain that $C$ is a feasible $k$-Path Cover

→ produces a minimal cover in the set theoretic sense!
→ specialized decision oracle for k-SPC

**Instance-based Lower Bounds**
→ number of pairwise disjoint k-paths



## EXPERIMENTAL RESULTS

**Soft/Hardware** C++, gcc 4.6.3  3.2GHz intel i5-3470, 16GB RAM

**Test Graphs**
Germany (GER): 18M nodes, 36M edges
USA: 24M nodes, 58M edges

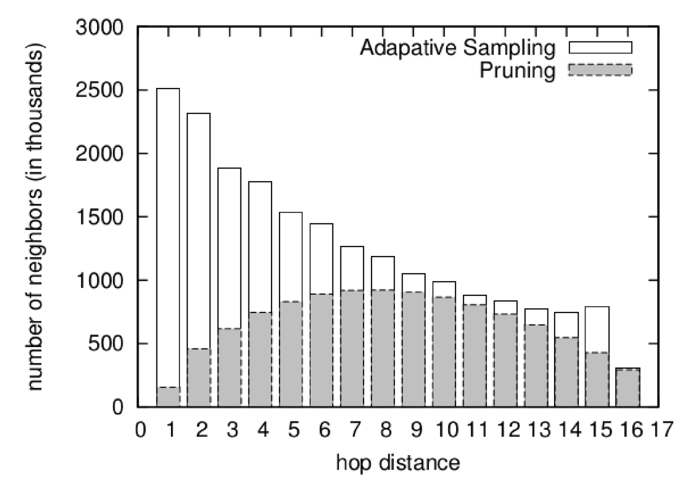
**k-APC**

| G | k | lower bound | \|C\| | perc | time(s) | apx |
|---|---|---|---|---|---|---|
| GER | 2 | 8,560,543 | 8,863,443 | 50.00% | 17 | 1.04 |
| | 4 | 3,969,092 | 4,513,217 | 25.50% | 21 | 1.14 |
| | 8 | 1,739,476 | 2,308,934 | 13.00% | 29 | 1.33 |
| | 16 | 735,746 | 1,209,215 | 6.82% | 47 | 1.64 |
| | 32 | 306,009 | 666,829 | 3.76% | 119 | 2.18 |
| USA | 2 | 10,906,996 | 11,910,322 | 49.70% | 15 | 1.09 |
| | 4 | 4,631,511 | 6,676,239 | 27.90% | 22 | 1.44 |
| | 8 | 1,854,605 | 3,776,360 | 15.80% | 38 | 2.04 |
| | 16 | 759,961 | 2,351,124 | 9.82% | 110 | 3.09 |
| | 32 | 321,853 | 1,603,267 | 6.69% | 15,100 | 4.98 |

+ results for varying node orders
→ very efficient computation up to $k = 16$
→ close-to-optimal results (provable via lower bounds)

**k-SPC**
Baseline: Adaptive Sampling (AS) [Tao et al., SIGMOD 2011]
cover sizes for $k = 16$ on USA
AS 14.0% of the nodes
Pruning 5.8% of the nodes

→ better compression of whole graph and single routes




### MAIN APPLICATION: PERSONALIZED ROUTE PLANNING

| k | Dijkstra (ms) | k-APC search (ms) | speed-up |
|---|---|---|---|
| 8 | 3,282 | 481 | 6.82 |
| 12 | 3,282 | 356 | 9.21 |
| 16 | 3,282 | 295 | 11.1 |
| 20 | 3,282 | 265 | 12.4 |
| 24 | 3,282 | 249 | 13.1 |
| 28 | 3,282 | 248 | 13.2 |

| Metrics |
|---|
| travel time |
| eucl. dist |
| height difference |
| energy |
| edge-type |
| speed |
| rand |
| unit |

→ experiments with up to 64 metrics
→ to-date fastest personalized route planning scheme