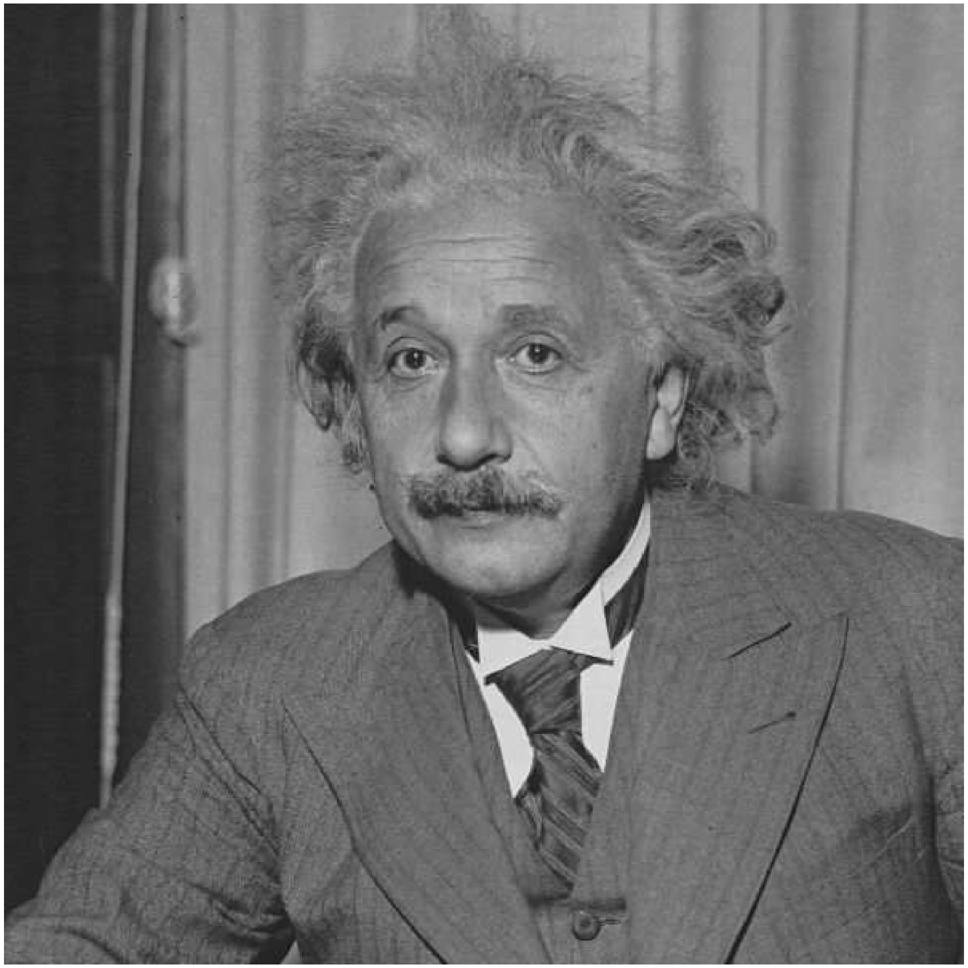
Point-to-Point Shortest Path Algorithms with Preprocessing

Andrew V. Goldberg
Microsoft Research – Silicon Valley
www.research.microsoft.com/∼goldberg/

Joint with Haim Kaplan and Renato Werneck

Einstein Quote



Everything should be made as simple as possible, but not simpler

Shortest Path Problem

Variants

- Nonnegative and arbitrary arc lengths.
- Point to point, single source, all pairs.
- Directed and undirected.

Here we study

- Point to point, nonnegative length, directed problem.
- Allow preprocessing with limited (linear) space.

Many applications, both directly and as a subroutine.

Shortest Path Problem

Input: Directed graph $G = (V, A)$, nonnegative length function $\ell : A \to \mathbf{R}^+$, origin $s \in V$, destination $t \in V$.

Preprocessing: Limited space to store results.

Query: Find a shortest path from s to t.

Interested in exact algorithms that search a (small) subgraph.

Related work: reach-based routing [Gutman 04], hierarchical decomposition [Schultz, Wagner & Weihe 02], [Sanders & Schultes 05, 06], geometric pruning [Wagner & Willhalm 03], arc flags [Lauther 04], [Köhler, Möhring & Schilling 05], [Möhring et al. 06].

Motivating Application

Driving directions

- Run on servers and small devices.
- Typical production codes
  - Use base graph or other heuristics based on road categories; needs hand-tuning.
  - Runs (bidirectional) Dijkstra or A* with Euclidean bounds on "patched" graph.
  - Non-exact and no performance guarantee.
- We are interested in exact and very efficient algorithms.
- New results finding their way into products.

Outline

- Scanning method and Dijkstra's algorithm.
- Bidirectional Dijkstra's algorithm.
- A^* search.
- ALT Algorithm
- Definition of reach
- Reach-based algorithm
- Reach for A^*

Scanning Method

- For each vertex v maintain its distance label $d_s(v)$ and status $S(v) \in \{\texttt{unreached}, \texttt{labeled}, \texttt{scanned}\}$.
- Unreached vertices have $d_s(v) = \infty$.
- If $d_s(v)$ decreases, v becomes labeled.
- To scan a labeled vertex v, for each arc (v, w), if $d_s(w) > d_s(v) + \ell(v, w)$ set $d_s(w) = d_s(v) + \ell(v, w)$.
- Initially for all vertices are unreached.
- Start by decreasing $d_s(s)$ to 0.
- While there are labeled vertices, pick one and scan it.
- Different selection rules lead to different algorithms.

Dijkstra's Algorithm

[Dijkstra 1959], [Dantzig 1963].

- At each step scan a labeled vertex with the minimum label.
- Stop when t is selected for scanning.

Work almost linear in the visited subgraph size.

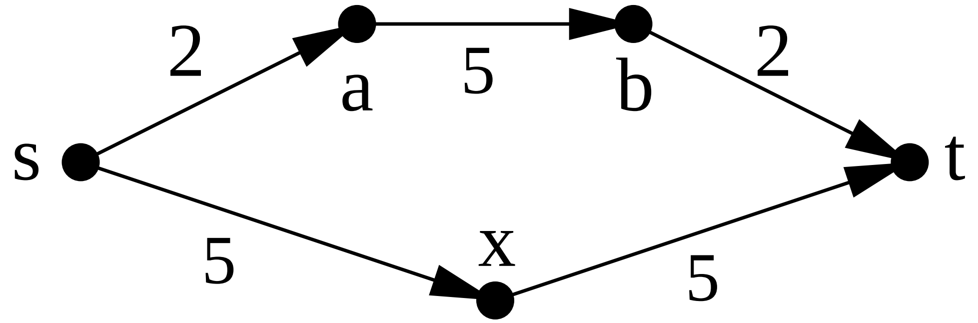
Reverse Algorithm: Run algorithm from t in the graph with all arcs reversed, stop when t is selected for scanning.

Bidirectional Algorithm

- Run forward Dijkstra from s and backward from t.
- Maintain μ, the length of the shortest path seen: when scanning an arc (v, w) such that w has been scanned in the other direction, check if the corresponding s-t path improves μ.
- Stop when about to scan a vertex x scanned in the other direction.
- Output μ and the corresponding path.

Bidirectional Algorithm: Pitfalls

The algorithm is not as simple as it looks.

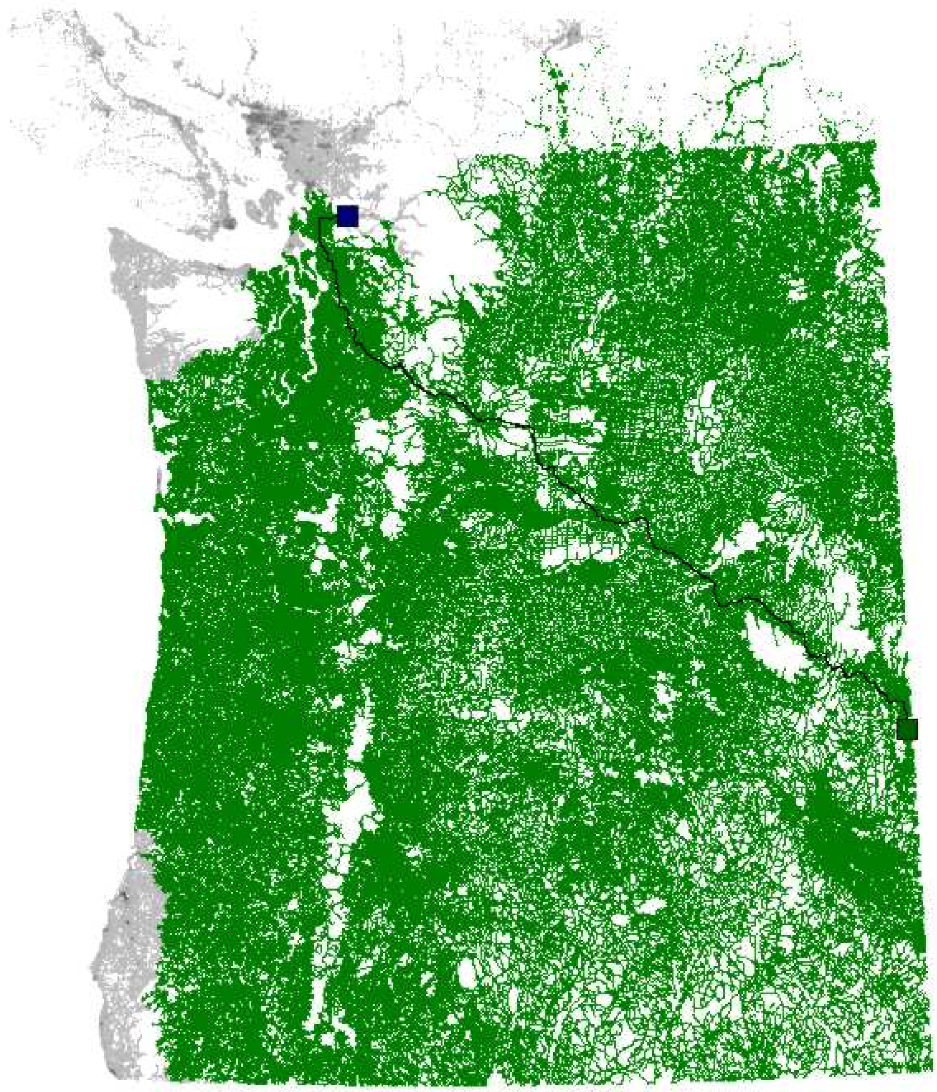


The searches meat at x, but x is not on the shortest path.

Example Graph

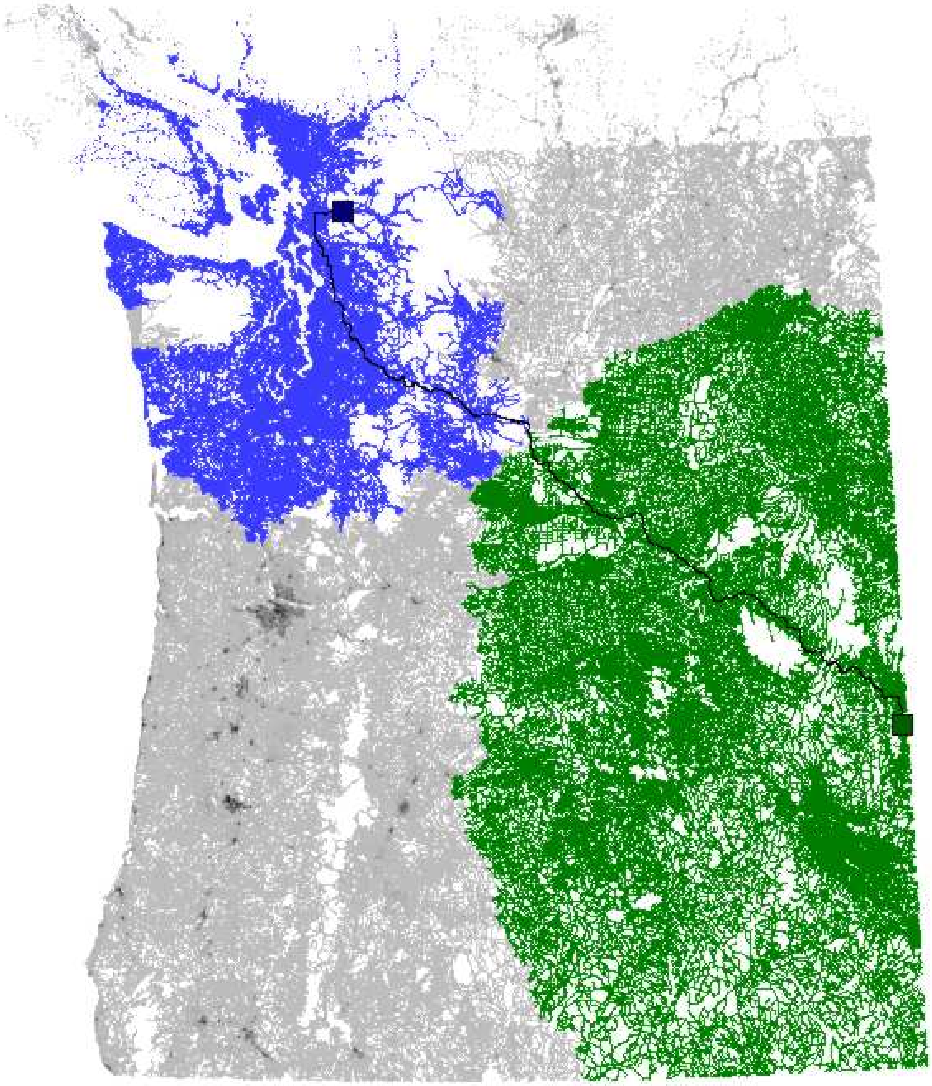
1.6M vertices, 3.8M arcs, travel time metric.

Dijkstra's Algorithm



Searched area

Bidirectional Algorithm



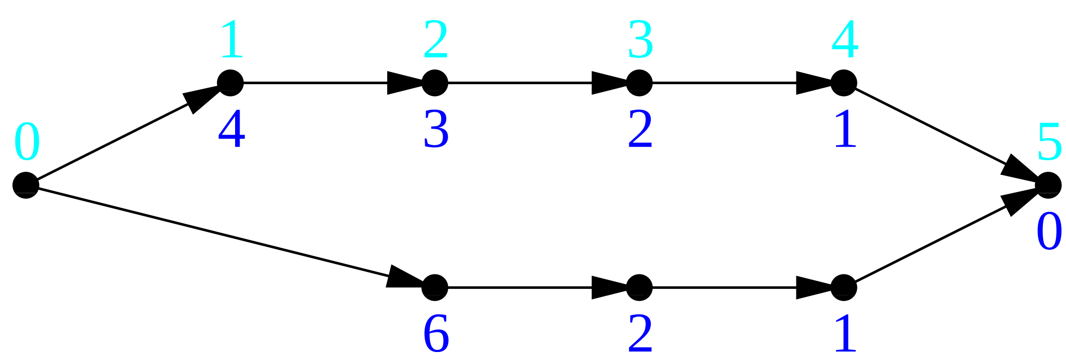
forward search/ reverse search

A* Search

[Doran 67], [Hart, Nilsson & Raphael 68]
Motivated by large search spaces (e.g., game graphs).

Similar to Dijkstra's algorithm but:

- Domain-specific estimates $\pi_t(v)$ on $\text{dist}(v, t)$ (potentials).
- At each step pick a labeled vertex with the minimum $k(v) = d_s(v) + \pi_t(v)$.
  Best estimate of path length.
- In general, optimality is not guaranteed.



Feasibility and Optimality

Potential transformation: Replace $\ell(v,w)$ by $\ell_{\pi_t}(v,w) = \ell(v,w) - \pi_t(v) + \pi_t(w)$ (reduced costs).

Fact: Problems defined by ℓ and ℓ_{π_t} are equivalent.

Definition: π_t is *feasible* if $\forall (v,w) \in A$, the reduced costs are nonnegative. (Estimates are "locally consistent".)

Optimality: If π_t is feasible, the A* search is equivalent to Dijkstra's algorithm on transformed network, which has nonnegative arc lengths. A* search finds an optimal path.

Different order of vertex scans, different subgraph searched.

Fact: If π_t is feasible and $\pi_t(t) = 0$, then π_t gives lower bounds on distances to t.

Computing Lower Bounds

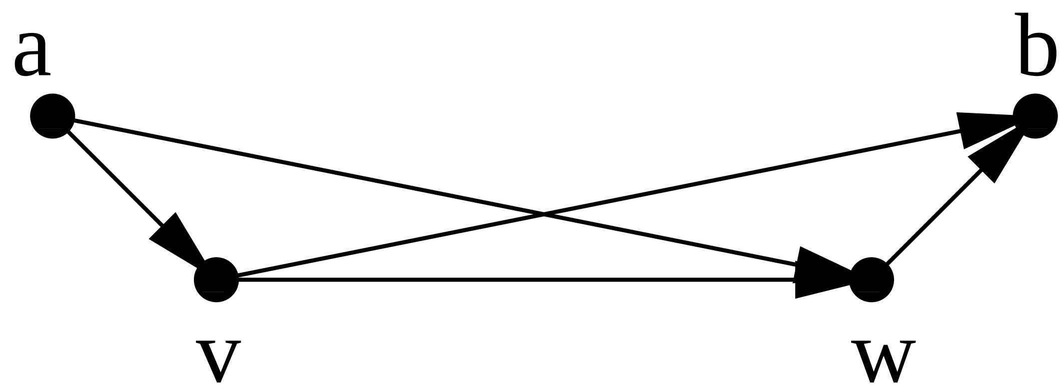
Euclidean bounds:

[folklore], [Pohl 71], [Sedgewick & Vitter 86].

For graph embedded in a metric space, use Euclidean distance.

Limited applicability, not very good for driving directions.

We use triangle inequality



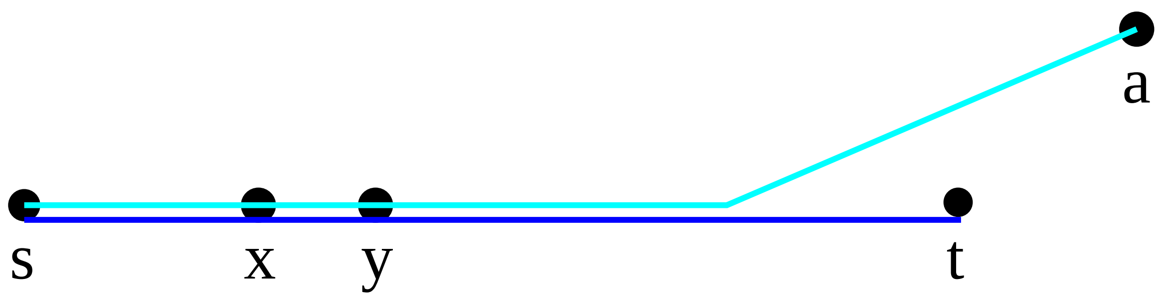
dist$(v, w) \geq$ dist$(v, b)-$dist(w, b); dist$(v, w) \geq$ dist$(a, w)-$dist(a, v).

Lower Bounds (cont.)

Maximum (minimum, average) of feasible potentials is feasible.

- Select landmarks (a small number).
- For all vertices, precompute distances to and from each landmark.
- For each s, t, use max of the corresponding lower bounds for $\pi_t(v)$.

Why this works well (when it does)



$\ell_{\pi_t}(x, y) = 0$

Bidirectional Lowerbounding

Forward reduced costs: $\ell_{\pi_t}(v, w) = \ell(v, w) - \pi_t(v) + \pi_t(w)$.

Reverse reduced costs: $\ell_{\pi_s}(v, w) = \ell(v, w) + \pi_s(v) - \pi_s(w)$.

What's the problem?

Bidirectional Lowerbounding

Forward reduced costs: $\ell_{\pi_t}(v,w) = \ell(v,w) - \pi_t(v) + \pi_t(w)$.

Reverse reduced costs: $\ell_{\pi_s}(v,w) = \ell(v,w) + \pi_s(v) - \pi_s(w)$.

Fact: π_t and π_s give the same reduced costs iff $\pi_s + \pi_t =$ const.

[Ikeda et at. 94]: use $p_s(v) = \frac{\pi_s(v) - \pi_t(v)}{2}$ and $p_t(v) = -p_s(v)$.

Other solutions possible. Easy to loose correctness.

ALT algorithms use A^* search and landmark-based lower bounds.

Landmark Selection

Preprocessing

- Random selection is fast.
- Many heuristics find better landmarks.
- Local search can find a good subset of candidate landmarks.
- We use a heuristic with local search.

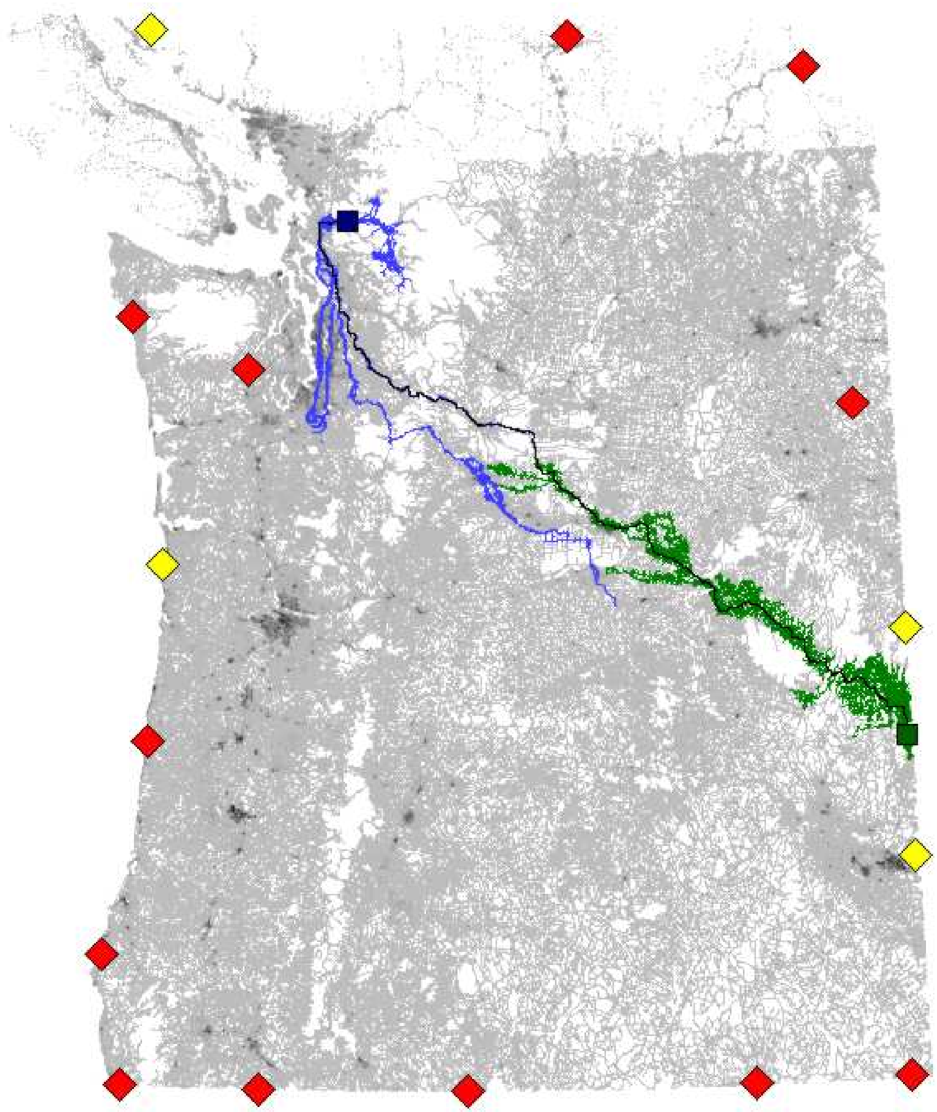
Preprocessing/query trade-off.

Query

- For a specific s, t pair, only some landmarks are useful.
- Use only active landmarks that give best bounds on dist(s, t).
- If needed, dynamically add active landmarks (good for the search frontier).

Allows using many landmarks with small time overhead.

Bidirectional ALT Example



Experimental Results

Northwest (1.6M vertices), random queries, 16 landmarks.

method	preprocessing		query		
	minutes	MB	avgscan	maxscan	ms
Bidirectional Dijkstra	—	28	518 723	1 197 607	340.74
ALT	4	132	16 276	150 389	12.05

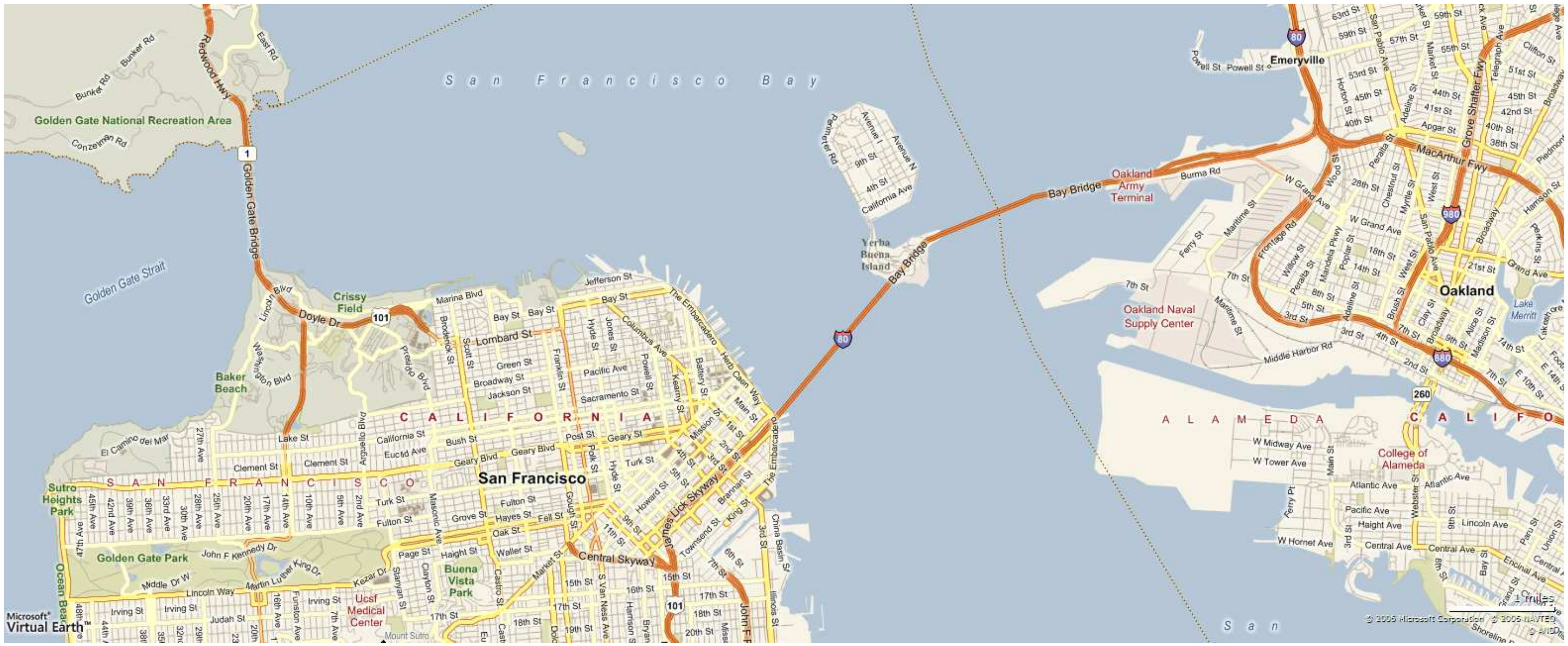
Related Systems Work

Network delay estimation:
Use delays to beacons to estimate arbitrary node delays.
E.g., IDMaps [Francis et al. 01].

Theoretical analysis [Kleinberg, Slivkins & Wexler 04]: for random beacons and bounded doubling dimension graphs, get good bounds for most node pairs.

Good bounds are not enough to prove bounds on ALT.

Reach Intuition



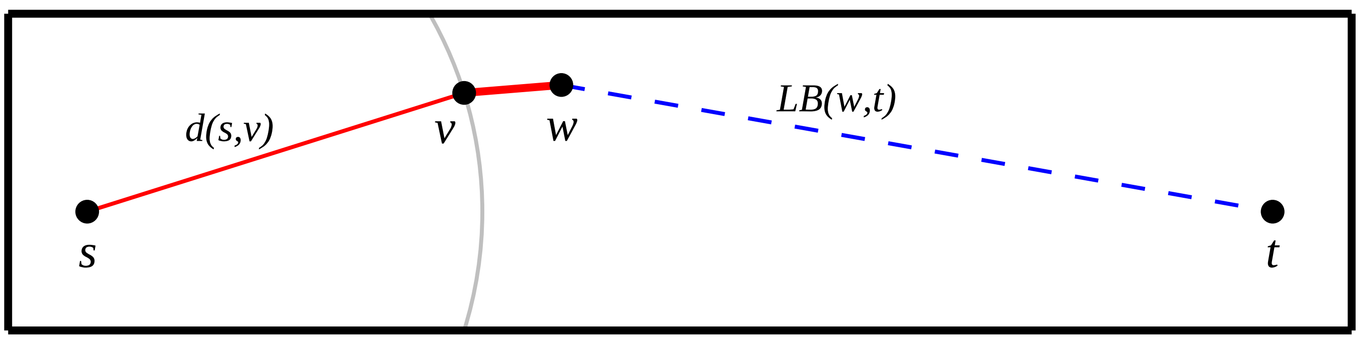
Identify local intersections and prune them when searching far from s and t.

Reaches

[Gutman 04]

- Consider a vertex v that splits a path P into P_1 and P_2. $r_P(v) = \min(\ell(P_1), \ell(P_2))$.
- $r(v) = \max_P(r_P(v))$ over all shortest paths P through v.

Using reaches to prune Dijkstra:

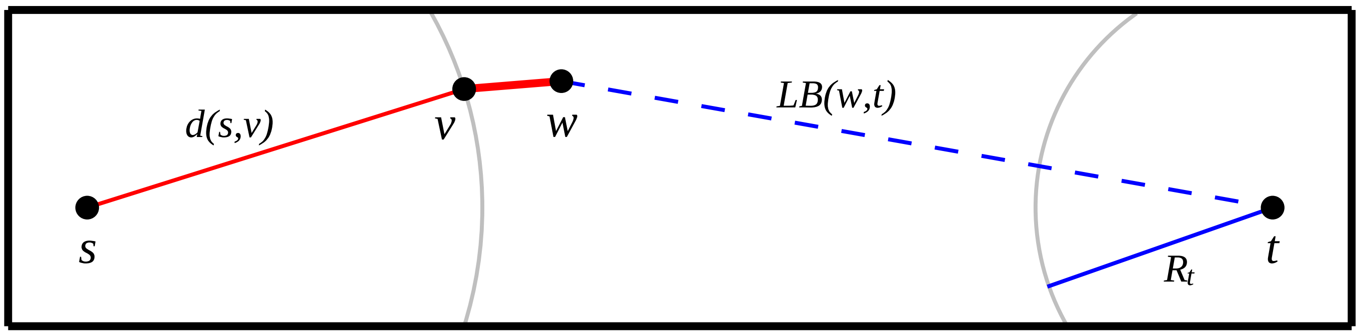


If $r(w) < \min(d(v) + \ell(v,w), LB(w,t))$ then prune w.

Obtaining Lower Bounds

Can use landmark lower bounds if available.

Bidirectional search gives implicit bounds (R_t below).



Reach-based query algorithm is Dijkstra's algorithm with pruning based on reaches. Given a lower-bound subroutine, a small change to Dijkstra's algorithm.

Computing Reaches

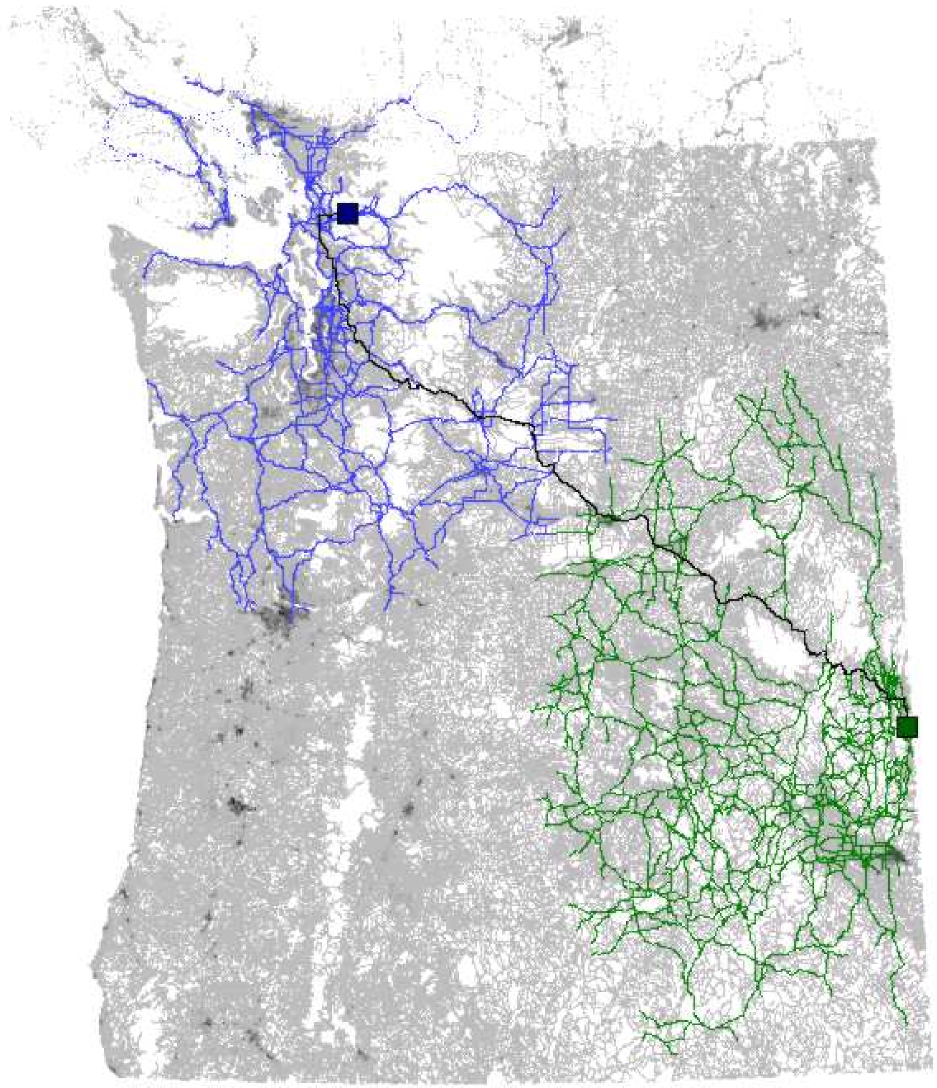
- A natural exact computation uses all-pairs shortest paths.
- Overnight for 0.3M vertex graph, years for 30M vertex graph.
- Have a heuristic improvement, but it is not fast enough.
- Can use reach upper bounds for query search pruning.

Iterative approximation algorithm: [Gutman 04]

- Use partial shortest path trees of depth $O(\epsilon)$ to bound reaches of vertices v with $r(v) < \epsilon$.
- Delete vertices with bounded reaches, add penalties.
- Increase ϵ and repeat.

Query time does not increase much; preprocessing faster but still not fast enough.

Reach Algorithm



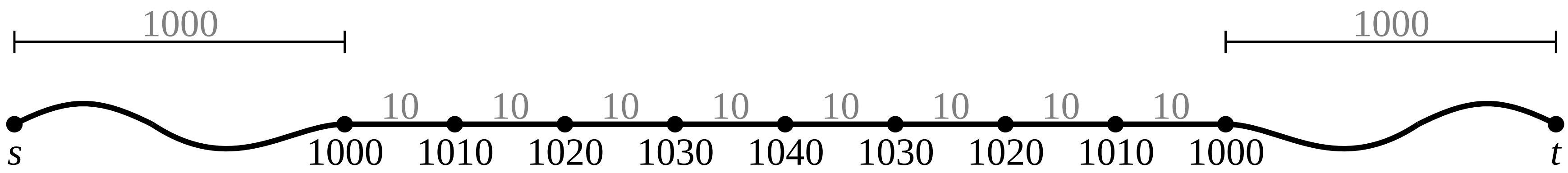
Experimental Results

Northwest (1.6M vertices), random queries, 16 landmarks.

method	preprocessing		query		
	minutes	MB	avgscan	maxscan	ms
Bidirectional Dijkstra	—	28	518 723	1 197 607	340.74
ALT	4	132	16 276	150 389	12.05
Reach	1 100	34	53 888	106 288	30.61

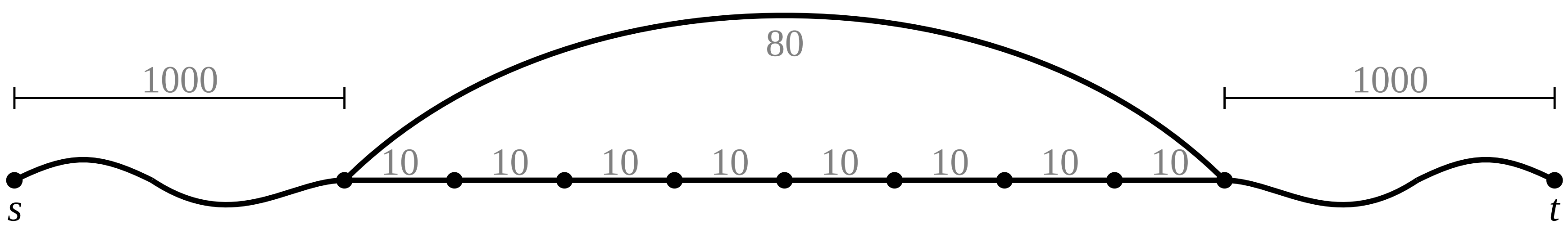
Shortcuts

- Consider the graph below.
- Many vertices have large reach.



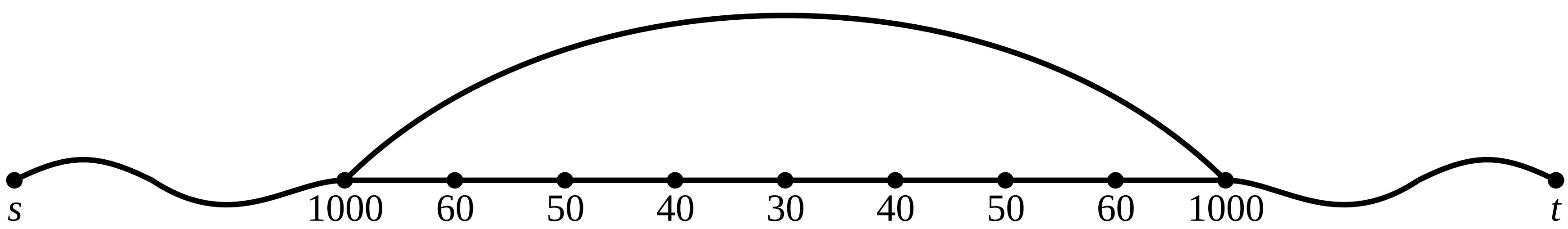
Shortcuts

- Consider the graph below.
- Many vertices have large reach.
- Add a shortcut arc, break ties by the number of hops.



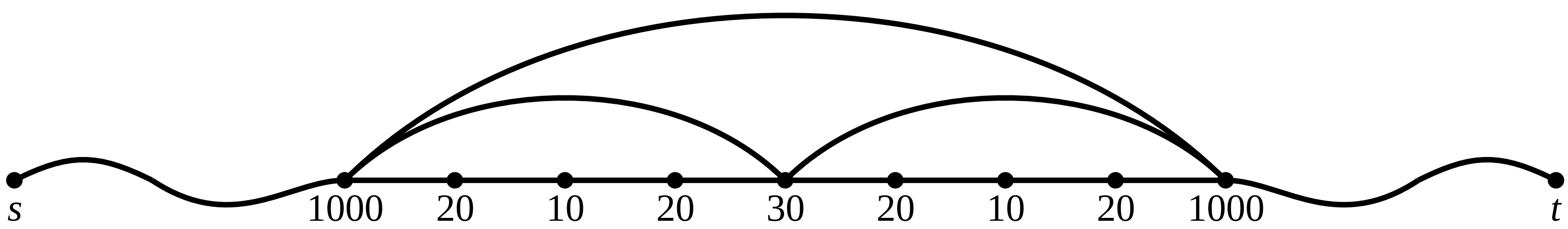
Shortcuts

- Consider the graph below.
- Many vertices have large reach.
- Add a shortcut arc, break ties by the number of hops.
- Reaches decrease.



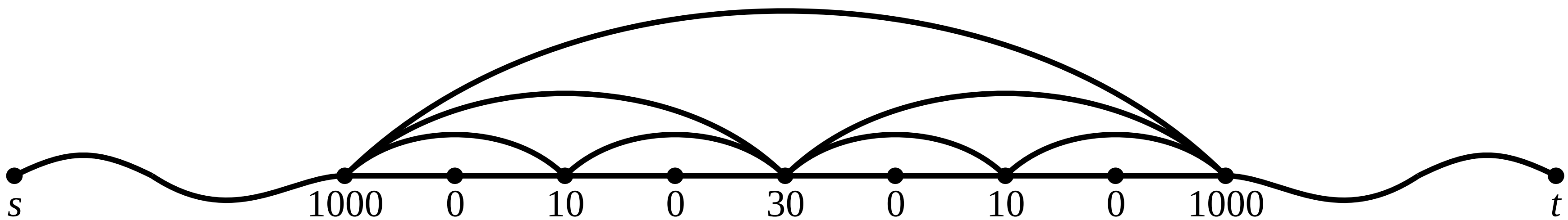
Shortcuts

- Consider the graph below.
- Many vertices have large reach.
- Add a shortcut arc, break ties by the number of hops.
- Reaches decrease.
- Repeat.



Shortcuts

- Consider the graph below.
- Many vertices have large reach.
- Add a shortcut arc, break ties by the number of hops.
- Reaches decrease.
- Repeat.
- A small number of shortcuts can greatly decrease many reaches.

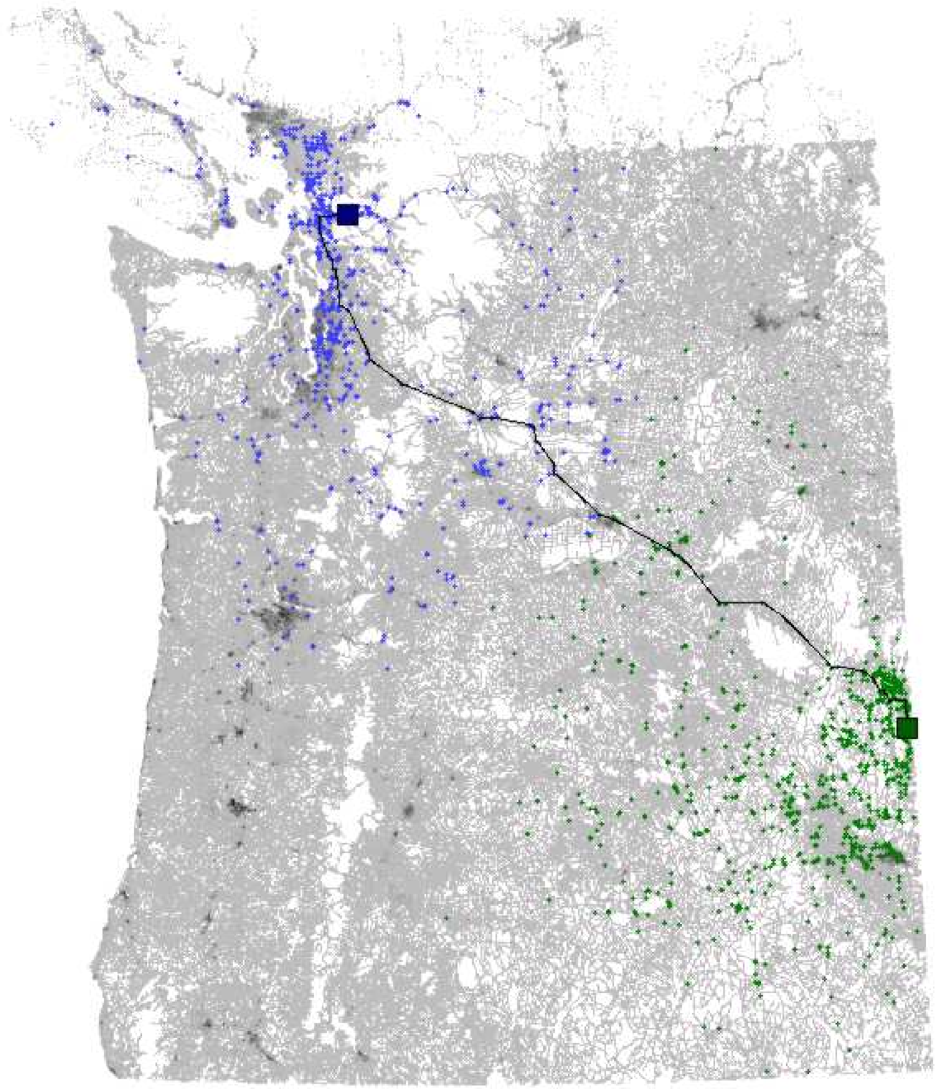


Shortcuts

[Sanders & Schultes 05, 06]: similar idea in hierarchy-based algorithm; similar performance.

- During preprocessing we shortcut small-degree vertices every time ϵ is updated.
- Shortcut replaces a vertex by a clique on its neighbors.
- A constant number of arcs is added for each deleted vertex.
- Shortcuts greatly speed up preprocessing.
- Shortcuts speed up queries.

Reach with Shortcuts



Experimental Results

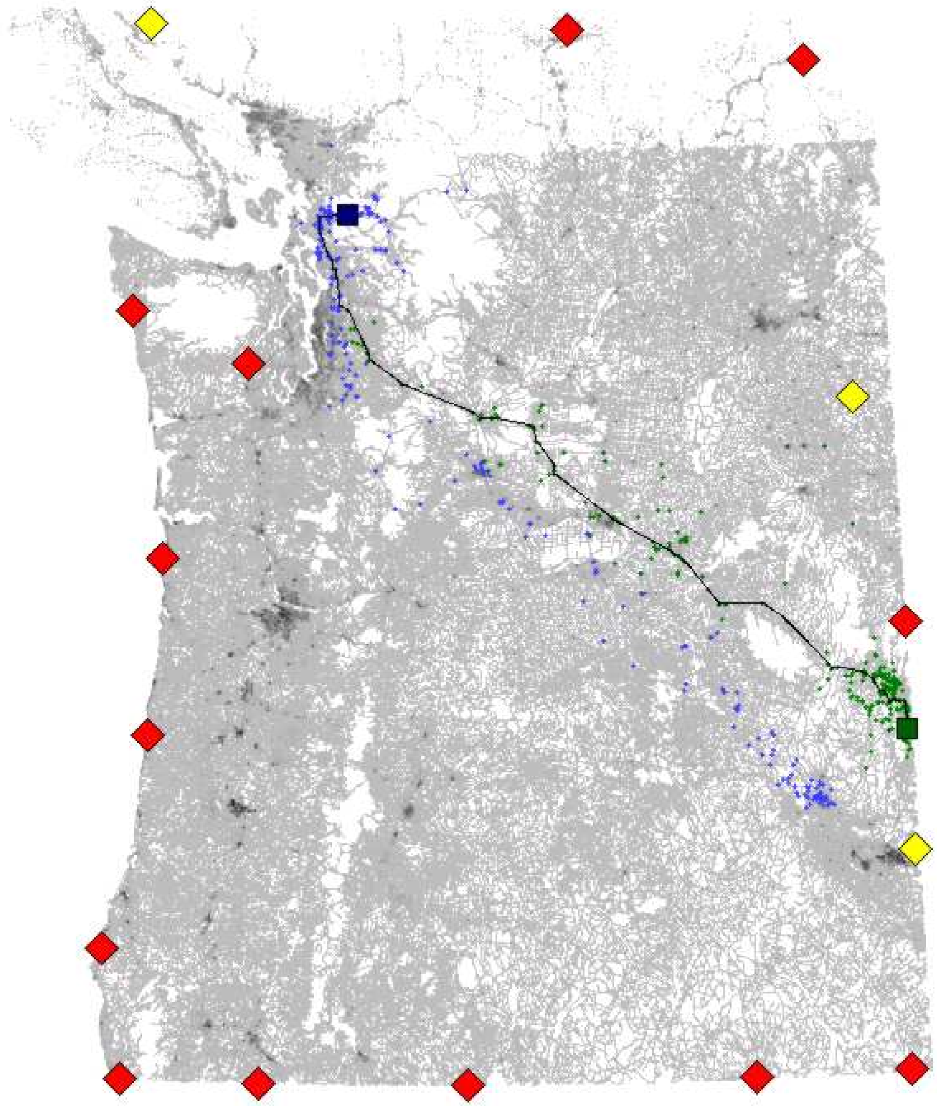
Northwest (1.6M vertices), random queries, 16 landmarks.

method	preprocessing		query		
	minutes	MB	avgscan	maxscan	ms
Bidirectional Dijkstra	—	28	518 723	1 197 607	340.74
ALT	4	132	16 276	150 389	12.05
Reach	1 100	34	53 888	106 288	30.61
Reach+Short	17	100	2 804	5 877	2.39

Reaches and ALT

- ALT computes transformed and original distances.
- ALT can be combined with reach pruning.
- Careful: Implicit lower bounds do not work, but landmark lower bounds do.
- Shortcuts do not affect landmark distances and bounds.

Reach with Shortcuts and ALT



Experimental Results

Northwest (1.6M vertices), random queries, 16 landmarks.

method	preprocessing		query		
	minutes	MB	avgscan	maxscan	ms
Bidirectional Dijkstra	—	28	518 723	1 197 607	340.74
ALT	4	132	16 276	150 389	12.05
Reach	1 100	34	53 888	106 288	30.61
Reach+Short	17	100	2 804	5 877	2.39
Reach+Short+ALT	21	204	367	1 513	0.73

Further Improvements

- Improved locality (sort by reach).
- For RE, factor of $3 - 12$ improvement for preprocessing and factor of $2 - 4$ for query times.
- Reach-aware landmarks: time/space trade-off.
- Idea: maintain landmark distances for a small fraction of high-reach vertices only.
- Can use more landmarks and improve both time and space.

Practical even for large (USA, Europe) graphs

- ≈ 1 ms. query time on a server.
- ≈ 5sec. query time on a Pocket PC with 2GB flash card.
- Better for local queries.

The USA Graph

USA: 24M vertices, 58M arcs, time metric, random queries.

method	preprocessing		query		
	min	KB	avgscan	maxscan	ms
Dijkstra	—	536	11 808 864	—	5 440.49
ALT(16)	17.6	2 563	187 968	2 183 718	295.44
Reach	impractical				
Reach+Short	27.9	893	2 405	4 813	1.77
Reach+Short+ALT(16,1)	45.5	3 032	592	2 668	0.80
Reach+Short+ALT(64,16)	113.9	1 579	538	2 534	0.86

The USA Graph

USA: 24M vertices, 58M arcs, distance metric, random queries.

method	preprocessing		query		
	min	KB	avgscan	maxscan	ms
Dijkstra	—	536	11 782 104	—	4 576.02
ALT(16)	15.2	2 417	276 195	2 910 133	410.73
Reach	impractical				
Reach+Short	46.4	918	7 311	13 886	5.78
Reach+Short+ALT(16,1)	61.5	2 923	905	5 510	1.41
Reach+Short+ALT(64,16)	120.5	1 575	670	3 499	1.22

Europe Graph

Europe: 18M vertices, 43M arcs, time metric, random queries.

method	preprocessing		query		
	min	KB	avgscan	maxscan	ms
Dijkstra	—	393	8 984 289	—	4 365.81
ALT(16)	12.5	1 597	82 348	993 015	120.09
Reach	impractical				
Reach+Short	45.1	648	4 371	8 486	3.06
Reach+Short+ALT(16,1)	57.7	1 869	714	3 387	0.89
Reach+Short+ALT(64,16)	102.6	1 037	610	2 998	0.91

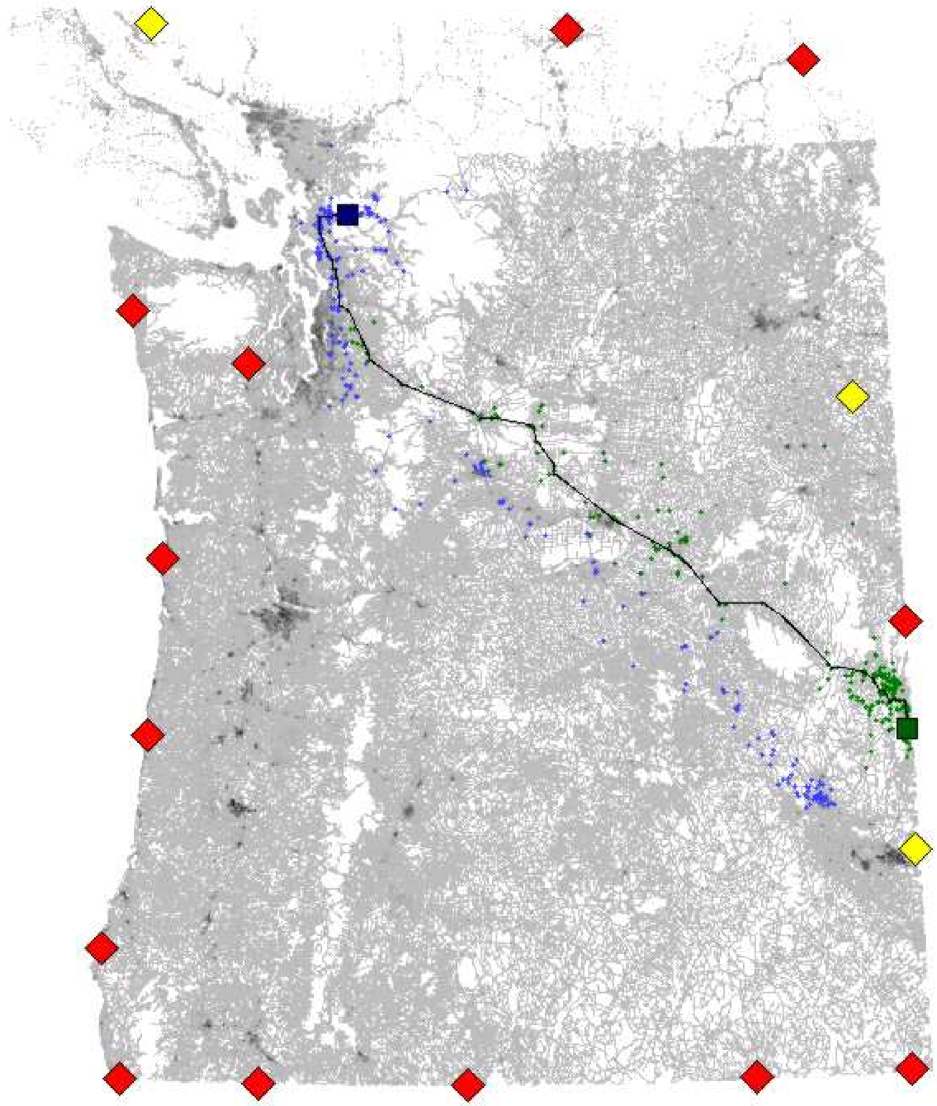
Grid Graphs

Grid with uniform random lengths (0.5M vertices), 16 landmarks.
No highway structure.

method	preprocessing		query		
	min	MB	avgscan	maxscan	ms
Bidirectional Dijkstra	—	18.0	174 150	416 925	160.14
ALT	0.26	96.6	6 057	65 664	6.28
Reach+Short	7.77	27.7	6 458	10 049	4.75
Reach+Short+ALT(16,1)	8.03	106.3	558	3 189	0.89
Reach+Short+ALT(64,16)	9.14	49.2	2 823	3 711	2.67

Reach preprocessing expensive, but helps queries.
(64,16) significantly slower that (16,1).

Demo



Concluding Remarks

- Our heuristics work well on road networks.
- Recent improvements: [Bast et al. 07, Geisberger et al. 08].
- How to select good shortcuts? (Road networks/grids.)
- For which classes of graphs do these techniques work?
- Need theoretical analysis for interesting graph classes.
- Interesting problems related to reach, e.g.
  - Is exact reach as hard as all-pairs shortest paths?
  - Constant-ratio upper bounds on reaches in $\tilde{O}(m)$ time.
- Dynamic graphs (real-time traffic).