Approximate Shortest Path and Distance Queries in Networks

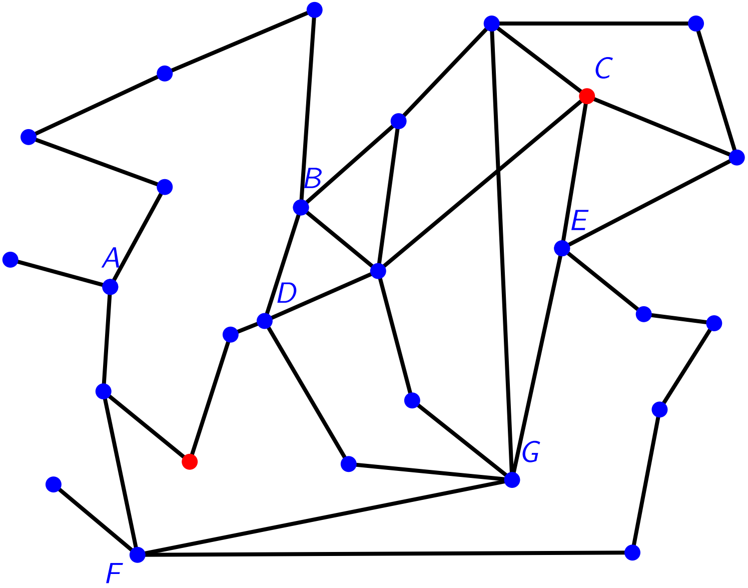
Christian Sommer

January 2010

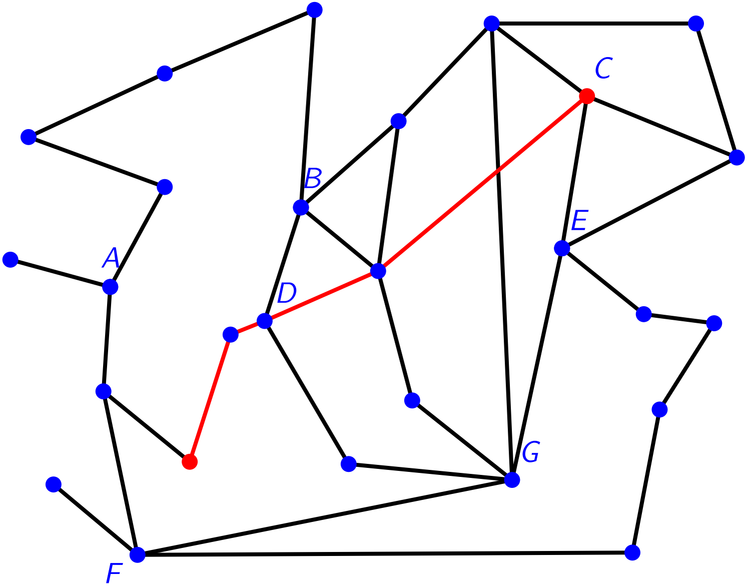
Outline

- Introduction
  - Motivation
  - Related Work
- Space Lower Bound
- Distances in Power-Law Graphs
- A Practical Method
- Conclusion

Motivation



Motivation



Motivation

Efficiently find a Shortest Path between Pairs of Nodes in

- Transportation Networks
- Social Networks
- Computer Networks (Internet)
- Protein Interaction Networks
- ...

Shortest Path Queries / Distance Oracles

- Preprocess a graph G with n nodes and m edges ...
- ... to create a Data Structure, using which ...
- ... we can efficiently answer Distance Queries.
  - $d(u, v)$

Shortest Path Queries / Distance Oracles

- Preprocess a graph G with n nodes and m edges ...
- ... to create a Data Structure, using which ...
- ... we can efficiently answer Approximate Distance Queries.
  - $\widetilde{d}(u, v)$

Approximate Distance Oracles — Stretch

- Distance between u and v in graph G: $d_G(u, v)$
- Oracle Result $\widetilde{d}(u, v)$ satisfies

$$d_G(u, v) \leqslant \widetilde{d}(u, v) \leqslant \alpha \cdot d_G(u, v) + \beta.$$

- Stretch (α, β)
  - Multiplicative Stretch α
  - Additive Stretch β

Related Work

Practical

- Focus on Transportation Networks

Theoretical

- General, undirected graphs
- Restricted classes (planar, bounded tree-width, minor-closed,...)

Related Work

Practical

- Focus on Transportation Networks

Theoretical

- General, undirected graphs
- Restricted classes (planar, bounded tree-width, minor-closed,...)

Shortest Path Queries in Transportation Networks

- Main focus, large body of research since 60's/70's
- Big progress around 2006 (DIMACS Implementation Challenge)
  - Preprocessing: tens of minutes for road map of the US/EU
  - Query time: $\approx 10^6$ times faster than Dijkstra's algorithm
- Ideas
  - Geometry, coordinates, A* search [SV86]
  - Goal-directed search (A* for graphs) [GH05]
  - Hierarchical structures [SS05, BFSS07, BD08, BDS$^+$08]
- Heuristics that work very well for road networks (often need separators)

Related Work

Practical

- Focus on Transportation Networks

Theoretical

- General, undirected graphs — unweighted if not stated otherwise
- Restricted classes (planar, bounded tree-width, minor-closed,...)

On Distance Oracles for General Undirected Graphs

Preprocessing	Space	Query	Stretch	Reference
$\mathcal{O}(mn)$	$\mathcal{O}(n^2)$	$\mathcal{O}(1)$	$(1, 0)$	APSP
$\mathcal{O}(1)$	$\mathcal{O}(m)$	$\mathcal{O}(m)$	$(1, 0)$	BFS
$\widetilde{\mathcal{O}}(kmn^{1/k})$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[TZ05, RTZ05]
$\mathcal{O}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[BS06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(n^{3/2})$	$\Theta(\lg n)$	$(3, 0)$	[BK06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[BK06], $k \geqslant 3$
$\mathcal{O}(m + n^{23/12})$	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(1)$	$(3, 10)$	[BGSU08]
$\widetilde{\mathcal{O}}(mn^{1/k})$	$\mathcal{O}(n^{1+1/k})$	$\mathcal{O}(1)$	$(\mathcal{O}(k), 0)$	[MN06, MS08]

On Distance Oracles for General Undirected Graphs

Preprocessing	Space	Query	Stretch	Reference
$\mathcal{O}(mn)$	$\mathcal{O}(n^2)$	$\mathcal{O}(1)$	$(1, 0)$	APSP
$\mathcal{O}(1)$	$\mathcal{O}(m)$	$\mathcal{O}(m)$	$(1, 0)$	BFS
$\widetilde{\mathcal{O}}(kmn^{1/k})$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[TZ05, RTZ05]
$\mathcal{O}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[BS06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(n^{3/2})$	$\Theta(\lg n)$	$(3, 0)$	[BK06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[BK06], $k \geqslant 3$
$\mathcal{O}(m + n^{23/12})$	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(1)$	$(3, 10)$	[BGSU08]
$\widetilde{\mathcal{O}}(mn^{1/k})$	$\mathcal{O}(n^{1+1/k})$	$\mathcal{O}(1)$	$(\mathcal{O}(k), 0)$	[MN06, MS08]

On Distance Oracles for General Undirected Graphs

Preprocessing	Space	Query	Stretch	Reference
$\mathcal{O}(mn)$	$\mathcal{O}(n^2)$	$\mathcal{O}(1)$	$(1,0)$	APSP
$\mathcal{O}(1)$	$\mathcal{O}(m)$	$\mathcal{O}(m)$	$(1,0)$	BFS
$\widetilde{\mathcal{O}}(kmn^{1/k})$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[TZ05, RTZ05]
$\mathcal{O}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[BS06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(n^{3/2})$	$\Theta(\lg n)$	$(3,0)$	[BK06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[BK06], $k \geqslant 3$
$\mathcal{O}(m+n^{23/12})$	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(1)$	$(3,10)$	[BGSU08]
$\widetilde{\mathcal{O}}(mn^{1/k})$	$\mathcal{O}(n^{1+1/k})$	$\mathcal{O}(1)$	$(\mathcal{O}(k),0)$	[MN06, MS08]
	$\Omega(n^{1+1/k})$		$< (2k+1,0)$	[TZ05]
	$n^{1+\Omega(1/t)}$	t	$< (1,2)$	thesis/[Pat08]
	$n^{1+\Omega(1/\alpha t)}$	t	$\leqslant (\alpha,0)$	this thesis

$\rightsquigarrow$ (Almost) Optimal Space/Stretch Tradeoff

On Distance Oracles for General Undirected Graphs

Preprocessing	Space	Query	Stretch	Reference
$\mathcal{O}(mn)$	$\mathcal{O}(n^2)$	$\mathcal{O}(1)$	$(1, 0)$	APSP
$\mathcal{O}(1)$	$\mathcal{O}(m)$	$\mathcal{O}(m)$	$(1, 0)$	BFS
$\widetilde{\mathcal{O}}(kmn^{1/k})$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k - 1, 0)$	[TZ05, RTZ05]
$\mathcal{O}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k - 1, 0)$	[BS06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(n^{3/2})$	$\Theta(\lg n)$	$(3, 0)$	[BK06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k - 1, 0)$	[BK06], $k \geqslant 3$
$\mathcal{O}(m + n^{23/12})$	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(1)$	$(3, 10)$	[BGSU08]
$\widetilde{\mathcal{O}}(mn^{1/k})$	$\mathcal{O}(n^{1+1/k})$	$\mathcal{O}(1)$	$(\mathcal{O}(k), 0)$	[MN06, MS08]

On Distance Oracles for General Undirected Graphs

Preprocessing	Space	Query	Stretch	Reference
$\mathcal{O}(mn)$	$\mathcal{O}(n^2)$	$\mathcal{O}(1)$	$(1, 0)$	APSP
$\mathcal{O}(1)$	$\mathcal{O}(m)$	$\mathcal{O}(m)$	$(1, 0)$	BFS
$\widetilde{\mathcal{O}}(kmn^{1/k})$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[TZ05, RTZ05]
$\mathcal{O}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[BS06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(n^{3/2})$	$\Theta(\lg n)$	$(3, 0)$	[BK06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[BK06], $k \geqslant 3$
$\mathcal{O}(m + n^{23/12})$	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(1)$	$(3, 10)$	[BGSU08]
$\widetilde{\mathcal{O}}(mn^{1/k})$	$\mathcal{O}(n^{1+1/k})$	$\mathcal{O}(1)$	$(\mathcal{O}(k), 0)$	[MN06, MS08]

Related Work

Practical

- Focus on Transportation Networks

Theoretical

- General, undirected graphs
- Restricted graph classes (planar, bounded tree-width, minor-closed,...)

On Distance Oracles for Special Classes of Graphs

- Bounded Tree-width [CZ00]
  Space $\mathcal{O}(n)$, Exact

- Planar [Tho04, Kle02] & Minor-free [AG06] & Bounded Doubling Dimension [Tal04]
  Space $\widetilde{\mathcal{O}}(n)$, Stretch $(1+\epsilon, 0)$

Outline

On Distance Oracles for General Undirected Graphs

Preprocessing	Space	Query	Stretch	Reference
$\mathcal{O}(mn)$	$\mathcal{O}(n^2)$	$\mathcal{O}(1)$	$(1,0)$	APSP
$\mathcal{O}(1)$	$\mathcal{O}(m)$	$\mathcal{O}(m)$	$(1,0)$	BFS
$\widetilde{\mathcal{O}}(kmn^{1/k})$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[TZ05, RTZ05]
$\mathcal{O}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[BS06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(n^{3/2})$	$\Theta(\lg n)$	$(3,0)$	[BK06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[BK06], $k \geqslant 3$
$\mathcal{O}(m+n^{23/12})$	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(1)$	$(3,10)$	[BGSU08]
$\widetilde{\mathcal{O}}(mn^{1/k})$	$\mathcal{O}(n^{1+1/k})$	$\mathcal{O}(1)$	$(\mathcal{O}(k),0)$	[MN06, MS08]
	$\Omega(n^{1+1/k})$		$< (2k+1,0)$	[TZ05]
	$n^{1+\Omega(1/t)}$	t	$< (1,2)$	thesis/[Pat08]
	$n^{1+\Omega(1/\alpha t)}$	t	$\leqslant (\alpha,0)$	this thesis

On Distance Oracles for Special Classes of Graphs

- Bounded Tree-width [CZ00]
  Space $\mathcal{O}(n)$, Exact
- Planar [Tho04, Kle02] & Minor-free [AG06] & Bounded Doubling Dimension [Tal04]
  Space $\widetilde{\mathcal{O}}(n)$, Stretch $(1+\epsilon, 0)$
- All Sparse Graphs?

Space Lower Bound

Theorem

- For sufficiently large graphs ($n \geqslant n_*$ nodes),
- any $(\alpha, 0)$–approximate distance oracle
- with query time at least t
- requires space

$$\mathcal{S} \geqslant n^{1+\Omega\left(\frac{1}{\alpha \mathrm{t}}\right)} / \lg n.$$

Comparison with Related Work

Preprocessing	Space	Query	Stretch	Reference
$\mathcal{O}(mn)$	$\mathcal{O}(n^2)$	$\mathcal{O}(1)$	$(1, 0)$	APSP
$\mathcal{O}(1)$	$\mathcal{O}(m)$	$\mathcal{O}(m)$	$(1, 0)$	BFS
$\widetilde{\mathcal{O}}(kmn^{1/k})$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[TZ05, RTZ05]
$\mathcal{O}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[BS06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(n^{3/2})$	$\Theta(\lg n)$	$(3, 0)$	[BK06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1, 0)$	[BK06], $k \geqslant 3$
$\mathcal{O}(m + n^{23/12})$	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(1)$	$(3, 10)$	[BGSU08]
$\widetilde{\mathcal{O}}(mn^{1/k})$	$\mathcal{O}(n^{1+1/k})$	$\mathcal{O}(1)$	$(\mathcal{O}(k), 0)$	[MN06, MS08]
	$\Omega(n^{1+1/k})$		$< (2k+1, 0)$	[TZ05]
	$n^{1+\Omega(1/t)}$	t	$< (1, 2)$	thesis/[Pat08]
	$n^{1+\Omega(1/\alpha t)}$	t	$\leqslant (\alpha, 0)$	this thesis

Comparison with Related Work — "Hard" Graphs

- "hard" graphs in Thorup and Zwick's lower bound have average degree $\Theta\left(n^{2/\alpha}\right)$ $\rightsquigarrow$ very dense
  TZ actually prove space $\Omega(m)$
- This proof: regular graphs with degree
  $\Theta\left(\left(\frac{8tw\alpha}{\lg n}\right)^{2/C}\right) \leqslant polylog(n)$
  - query time t
  - word length w
  - (multiplicative) stretch α
  - constant $C \in [0, 1]$

  proves space $\omega(m)$

- Distance oracles for sparse graphs also require large space:
  Before: $\Omega(n \cdot polylog(n))$
  Now: $\Omega\left(n^{1+\epsilon}\right)$

Proof Idea

- Reduction from Distance Oracle to a Communication Protocol
- This Protocol efficiently solves the LopsidedSetDisjointness Problem from Communication Complexity
- But: There is a Communication Lower Bound for the LopsidedSetDisjointness Problem
- $\rightsquigarrow$ Space Lower Bound for the Distance Oracle

LopsidedSetDisjointness

Alice		Bob
$S_{\text{Alice}} \subseteq \mathcal{U}$	$\mathcal{U}$	$S_{\text{Bob}} \subseteq \mathcal{U}$
	$\longrightarrow$	
	$\longleftarrow$	
	$\longrightarrow$	
	$\longleftarrow$	

$$S_{\text{Alice}} \cap S_{\text{Bob}} \stackrel{?}{=} \emptyset$$

Data Structure reduces to Communication Protocol

Alice		Bob
$S_{\text{Alice}} \subseteq \mathcal{U}$	$\mathcal{U}$	$S_{\text{Bob}} \subseteq \mathcal{U}$
$f(S_{\text{Alice}})$	$f : \mathcal{U} \leftrightarrow E$	$f(S_{\text{Bob}})$
Query $\uparrow$		Data Structure $\uparrow$
t rounds		size $\mathcal{S}$
	$\xrightarrow{\lg \mathcal{S}}$	word length w
	$\xleftarrow{w}$	
	$\longrightarrow$	
	$\longleftarrow$	

$$S_{\text{Alice}} \cap S_{\text{Bob}} \stackrel{?}{=} \emptyset$$

[PT06, Pat08]

"Hard" Graphs

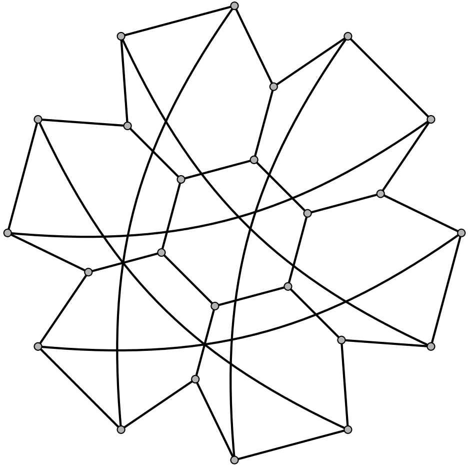
Expander Graphs [LPS88]

- n nodes, degree $\Theta\left(\left(\frac{8tw\alpha}{\lg n}\right)^{2/C}\right) \leqslant polylog(n)$
- girth $\mathcal{O}(\lg n)$ (girth: length of shortest cycle)
- many disjoint paths ($\approx \alpha$ times shorter than the girth)

- Distance oracle for sparse graphs must be able to handle expander graphs and all their subgraphs
- $(\alpha, 0)$–approximate distance oracle must handle distance queries in time t
  - (in particular for the endpoints of these paths ↑)

Almost LOPSIDEDSETDISJOINTNESS

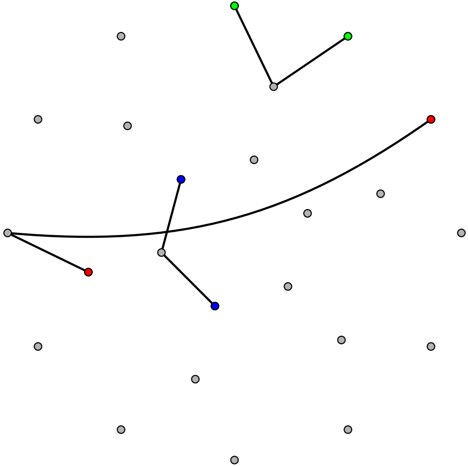
Alice	Bob



$G = (V, E)$

Almost LopsidedSetDisjointness

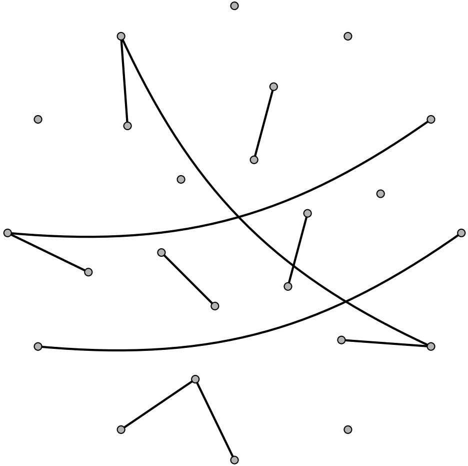
Alice	Bob



$f(S_{\text{Alice}})$

Almost LopsidedSetDisjointness

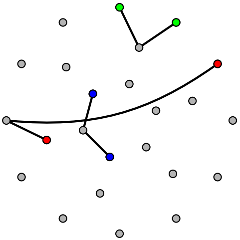
Alice	Bob

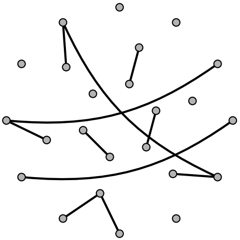


$f(S_{\text{Bob}})$

Almost LopsidedSetDisjointness

Alice		Bob
$f(S_{\text{Alice}})$		$f(S_{\text{Bob}})$





t rounds

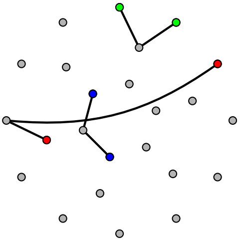
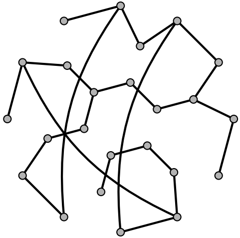
size $\mathcal{S}$

$\xrightarrow{\lg \mathcal{S}}$

word length w

$\xleftarrow{w}$

Almost LopsidedSetDisjointness

Alice	Bob
$f(S_{\text{Alice}})$	$G' = (V, E \setminus f(S_{\text{Bob}}))$
	
t rounds	Oracle for G', size $\mathcal{S}$
	word length w

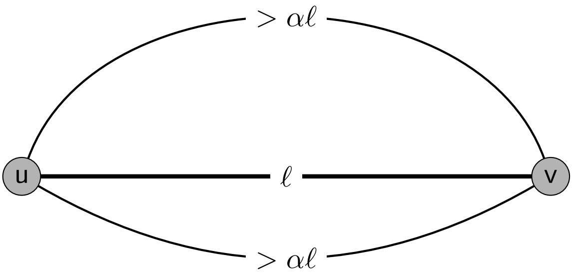
$\xrightarrow{\lg \mathcal{S}}$

$\xleftarrow{w}$

$S_{\text{Alice}} \cap S_{\text{Bob}} \stackrel{?}{=} \emptyset$

Path Queries and the Girth – not "too long" paths

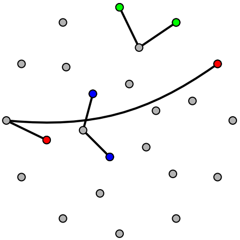
- query $\widetilde{d}_{G'}(u, v)$ for a shortest path from u to v in G (path not "too long")
  $\rightsquigarrow$ decide whether the path is contained in G'

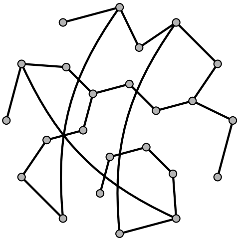


- $\widetilde{d}_{G'}(u, v) \leqslant \alpha\ell \rightsquigarrow$ all edges of the path are in $E(G')$
- $\widetilde{d}_{G'}(u, v) > \alpha\ell$ otherwise (since the girth is large)

Almost LOPSIDEDSETDISJOINTNESS

Alice	Bob
$f(S_{\text{Alice}})$	$G' = (V, E \setminus f(S_{\text{Bob}}))$





t rounds	$\xrightarrow{\lg \mathcal{S}}$	Oracle for G', size $\mathcal{S}$
	$\xleftarrow{w}$	word length w

$S_{\text{Alice}} \cap S_{\text{Bob}} \stackrel{?}{=} \emptyset$

Contribution of this Chapter

Preprocessing	Space	Query	Stretch	Reference
$\mathcal{O}(mn)$	$\mathcal{O}(n^2)$	$\mathcal{O}(1)$	$(1,0)$	APSP
$\mathcal{O}(1)$	$\mathcal{O}(m)$	$\mathcal{O}(m)$	$(1,0)$	BFS
$\widetilde{\mathcal{O}}(kmn^{1/k})$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[TZ05, RTZ05]
$\mathcal{O}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[BS06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(n^{3/2})$	$\Theta(\lg n)$	$(3,0)$	[BK06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[BK06], $k \geqslant 3$
$\mathcal{O}(m+n^{23/12})$	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(1)$	$(3,10)$	[BGSU08]
$\widetilde{\mathcal{O}}(mn^{1/k})$	$\mathcal{O}(n^{1+1/k})$	$\mathcal{O}(1)$	$(\mathcal{O}(k),0)$	[MN06, MS08]
	$\Omega(n^{1+1/k})$		$< (2k+1,0)$	[TZ05]
	$n^{1+\Omega(1/t)}$	t	$< (1,2)$	thesis/[Pat08]
	$n^{1+\Omega(1/\alpha t)}$	t	$\leqslant (\alpha,0)$	this thesis

Outline

- Introduction
- Space Lower Bound
- Distances in Power-Law Graphs
- A Practical Method
- Conclusion

On Distance Oracles for General Undirected Graphs

Preprocessing	Space	Query	Stretch	Reference
$\mathcal{O}(mn)$	$\mathcal{O}(n^2)$	$\mathcal{O}(1)$	$(1,0)$	APSP
$\mathcal{O}(1)$	$\mathcal{O}(m)$	$\mathcal{O}(m)$	$(1,0)$	BFS
$\widetilde{\mathcal{O}}(kmn^{1/k})$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[TZ05, RTZ05]
$\mathcal{O}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[BS06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(n^{3/2})$	$\Theta(\lg n)$	$(3,0)$	[BK06]
$\widetilde{\mathcal{O}}(n^2)$	$\mathcal{O}(kn^{1+1/k})$	$\mathcal{O}(k)$	$(2k-1,0)$	[BK06], $k \geqslant 3$
$\mathcal{O}(m+n^{23/12})$	$\mathcal{O}(n^{3/2})$	$\mathcal{O}(1)$	$(3,10)$	[BGSU08]
$\widetilde{\mathcal{O}}(mn^{1/k})$	$\mathcal{O}(n^{1+1/k})$	$\mathcal{O}(1)$	$(\mathcal{O}(k),0)$	[MN06, MS08]
	$\Omega(n^{1+1/k})$		$< (2k+1,0)$	[TZ05]
	$n^{1+\Omega(1/t)}$	t	$< (1,2)$	thesis/[Pat08]
	$n^{1+\Omega(1/\alpha t)}$	t	$\leqslant (\alpha,0)$	this thesis

On Distance Oracles for Special Classes of Graphs

- Bounded Tree-width [CZ00]
  Space $\mathcal{O}(n)$, Exact

- Planar [Tho04, Kle02] & Minor-free [AG06] & Bounded Doubling Dimension [Tal04]
  Space $\widetilde{\mathcal{O}}(n)$, Stretch $(1+\epsilon, 0)$

- Road Networks
  Efficient in Practice

- Other Real-World Networks?

Power-law Graphs

Power Law Distribution

- Random Variable D, Power Law Exponent $\tau \in (2,3)$

$$\Pr[D = x] \sim x^{-\tau}$$

Many Complex Networks are reported to have Power-law Degree Sequences $\rightsquigarrow$ most nodes have few neighbors, few nodes have high degrees

- Social Networks
- Internet (?)
- Protein Interaction Networks
- Citation Graph
- Web Graph (Hyperlinks)
- ...

Distance Oracles for Power-law Graphs?

- Unbounded Tree-width
- Not Planar & Not Minor-free & Unbounded Doubling Dimension
  (No small separators)
- $\rightsquigarrow$ General Distance Oracle?
  - Experiments show: [TZ05] works quite well (small space consumption) [KFY04, Section IV.B]

Distance Oracles for Power-law Graphs

- Experiments show: [TZ05] works quite well (small space consumption) [KFY04, Section IV.B]
- Common heuristic: route through nodes with high degree

- This thesis: proof why [TZ05] works well on Power-law Graphs and why high-degree heuristic is good
  - Chung-Lu Random Power-law Graphs
  - [TZ05] using high-degree nodes uses space $\mathcal{O}(n^{1+\gamma})$, $\gamma < \frac{1}{3}$ with high probability
    instead of space $\mathcal{O}(n^{1+\frac{1}{2}})$

Distance Oracle by Thorup & Zwick

Preprocessing

- Random Sampling of $\mathcal{O}(\sqrt{n})$ Landmarks
- SSSP for each Landmark
- Ball $B(u)$ for each Node u (expected size $\mathcal{O}(\sqrt{n})$)

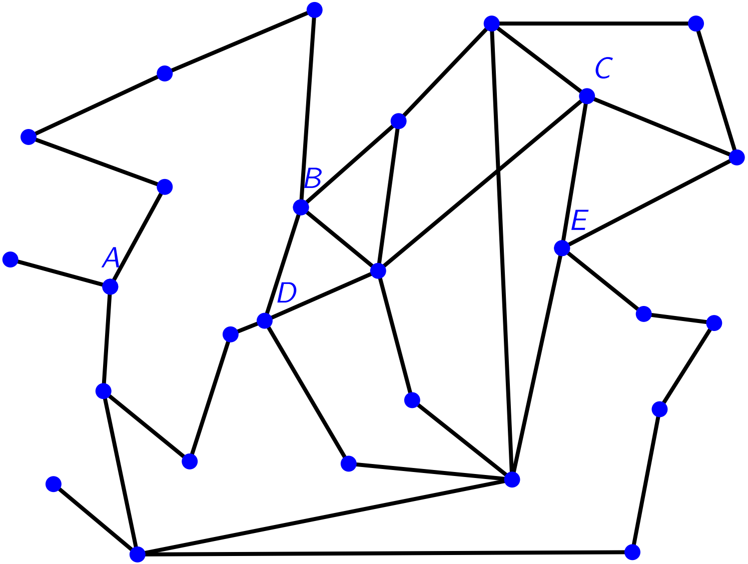
$$B(u) := \{v \in V(G) : d_G(u,v) < d_G(u,\ell_u)\},$$

  where ℓ_u denotes u's nearest landmark

Query $d(u,v)$

- If $v \in B(u)$ (or vice versa $u \in B(v)$) return exact distance
- Otherwise $u \to \ell_v \to v$, where ℓ_v denotes v's nearest landmark

Random Sampling



Distance Oracle by Thorup & Zwick

Preprocessing

- Random Sampling of $\mathcal{O}(\sqrt{n})$ Landmarks
- SSSP for each Landmark
- Ball $B(u)$ for each Node u (expected size $\mathcal{O}(\sqrt{n})$)

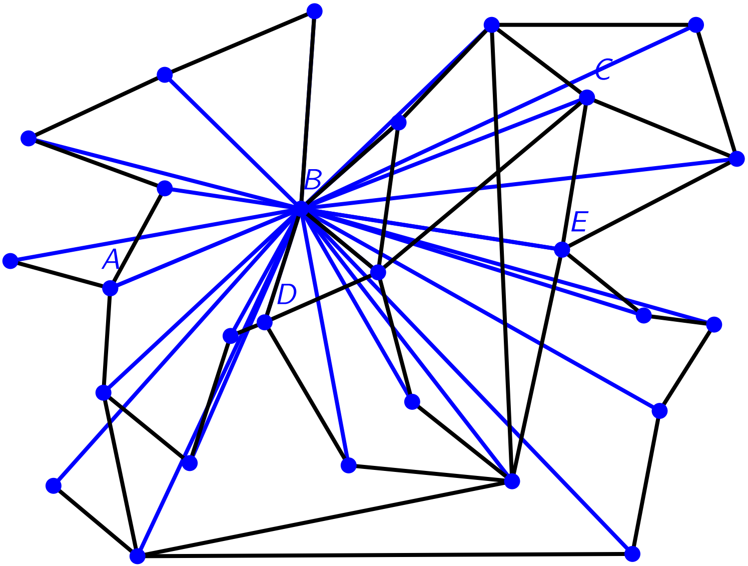
$$B(u) := \{v \in V(G) : d_G(u, v) < d_G(u, \ell_u)\},$$

where ℓ_u denotes u's nearest landmark

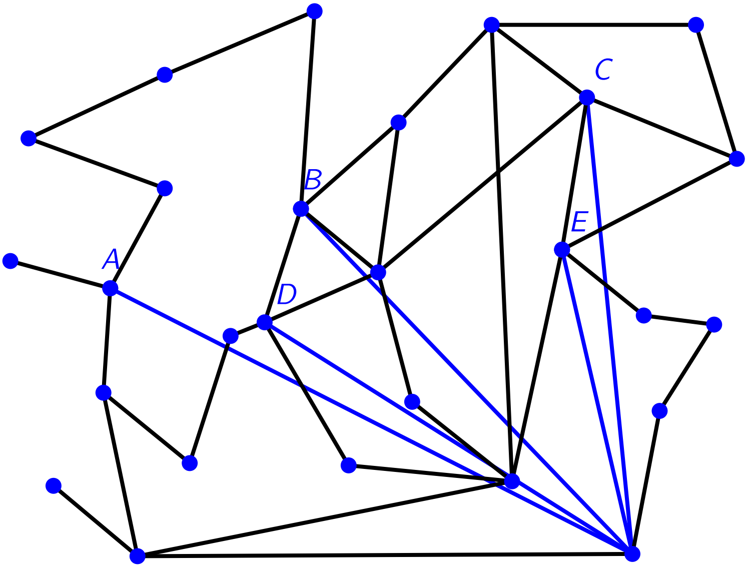
Query $d(u, v)$

- If $v \in B(u)$ (or vice versa $u \in B(v)$) return exact distance
- Otherwise $u \to \ell_v \to v$, where ℓ_v denotes v's nearest landmark

SSSP for all landmarks (fig. for one landmark)



SSSP for all landmarks (fig. for one node)



Distance Oracle by Thorup & Zwick

Preprocessing

- Random Sampling of $\mathcal{O}(\sqrt{n})$ Landmarks
- SSSP for each Landmark
- Ball $B(u)$ for each Node u (expected size $\mathcal{O}(\sqrt{n})$)

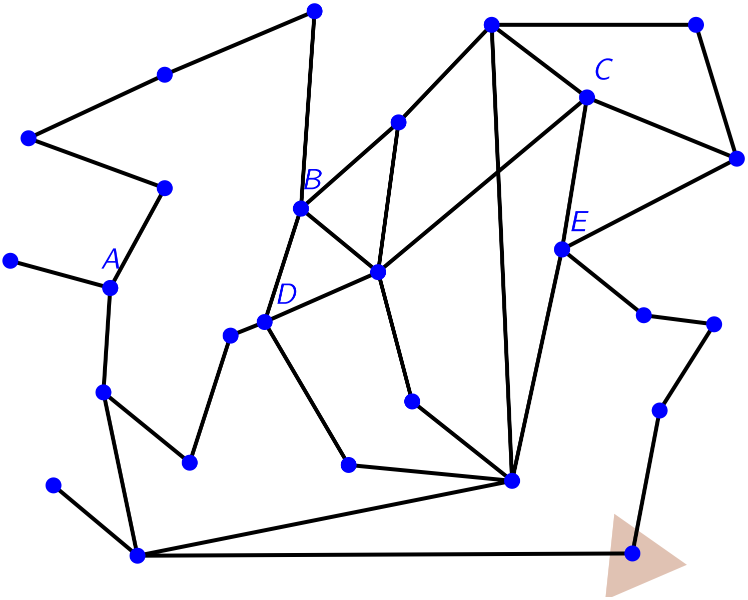
$$B(u) := \{v \in V(G) : d_G(u, v) < d_G(u, \ell_u)\},$$

where ℓ_u denotes u's nearest landmark

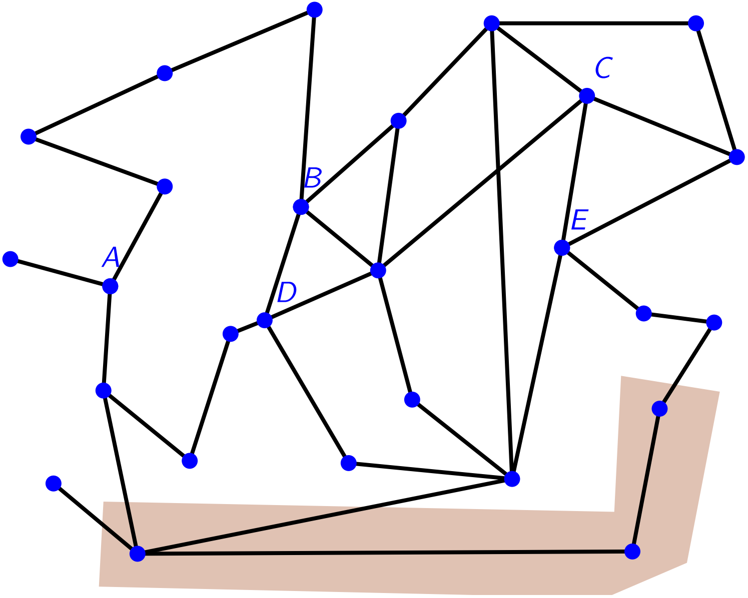
Query $d(u, v)$

- If $v \in B(u)$ (or vice versa $u \in B(v)$) return exact distance
- Otherwise $u \to \ell_v \to v$, where ℓ_v denotes v's nearest landmark

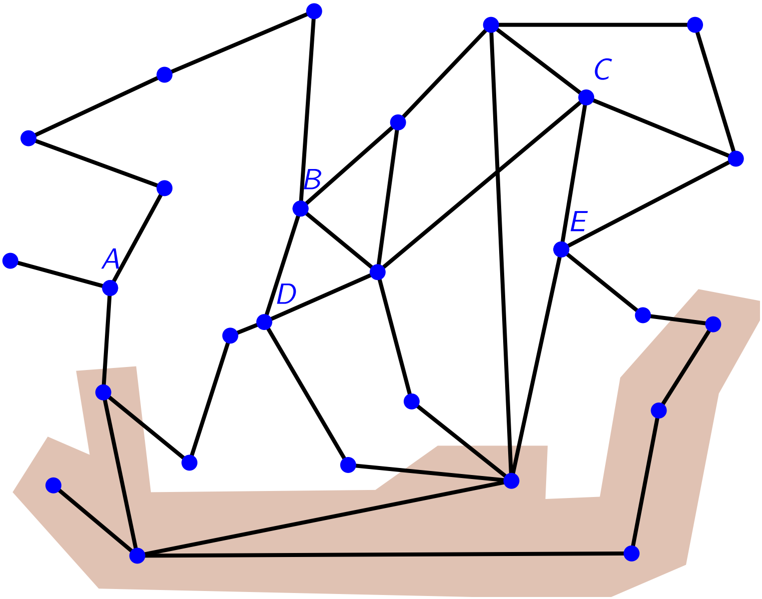
Ball Computation (fig. for one node)



Ball Computation (fig. for one node)



Ball Computation (fig. for one node)



Distance Oracle by Thorup & Zwick

Preprocessing

- Random Sampling of $\mathcal{O}(\sqrt{n})$ Landmarks
- SSSP for each Landmark
- Ball $B(u)$ for each Node u (expected size $\mathcal{O}(\sqrt{n})$)

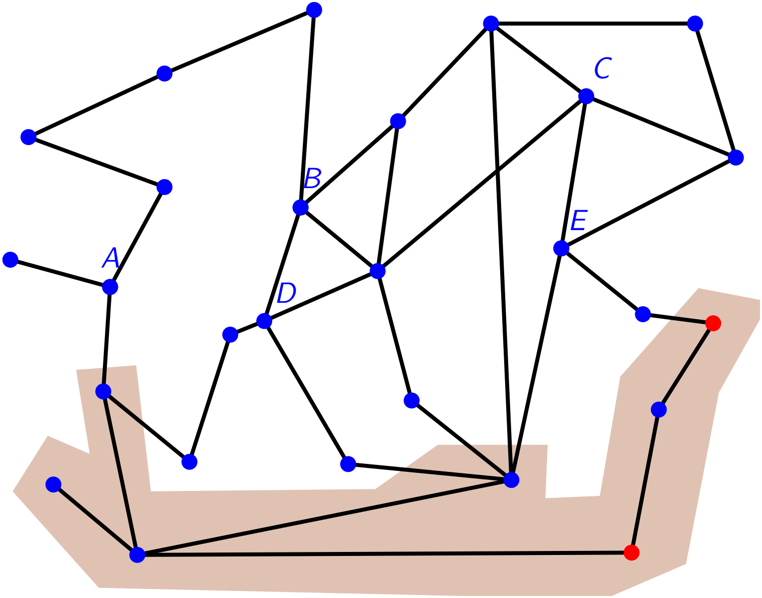
$$B(u) := \{v \in V(G) : d_G(u,v) < d_G(u,\ell_u)\},$$

  where ℓ_u denotes u's nearest landmark

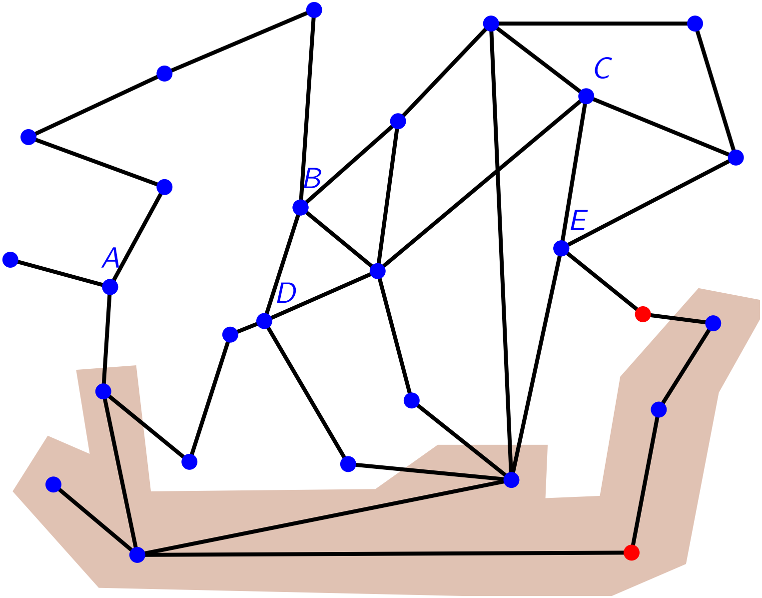
Query $d(u,v)$

- If $v \in B(u)$ (or vice versa $u \in B(v)$) return exact distance
- Otherwise $u \to \ell_v \to v$, where ℓ_v denotes v's nearest landmark

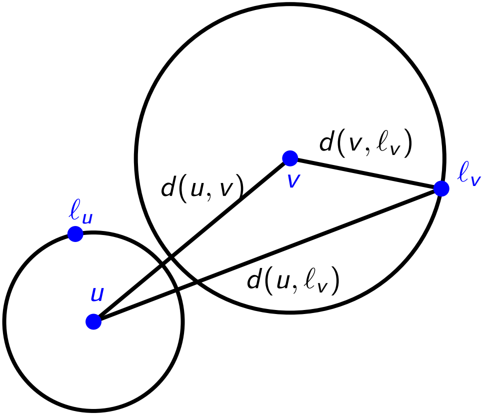
Query Source in Ball



Query Source not in Ball

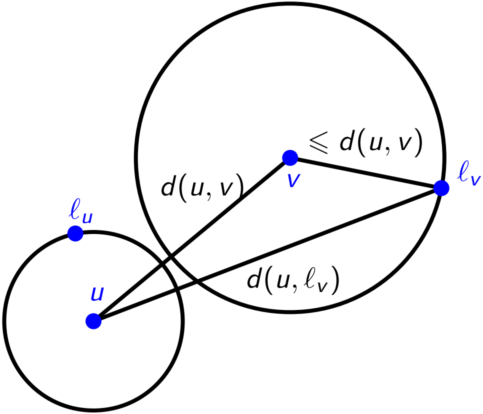


Multiplicative Stretch 3



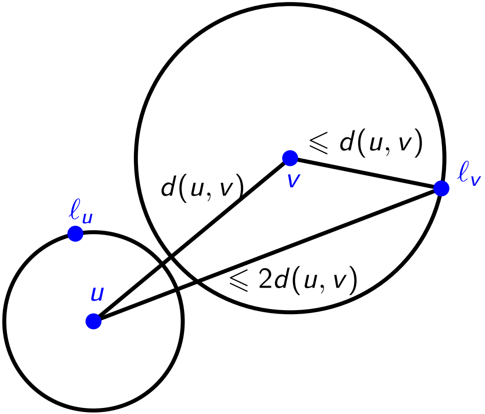
$$d(u, \ell_v) + d(v, \ell_v) \leqslant 3d(u, v)$$

Multiplicative Stretch 3



$$d(u, \ell_v) + d(v, \ell_v) \leqslant 3d(u, v)$$

Multiplicative Stretch 3



$$d(u,\ell_v) + d(v,\ell_v) \leqslant 3d(u,v)$$

Distance Oracle by Thorup & Zwick

Preprocessing

- Random Sampling of $\mathcal{O}(\sqrt{n})$ Landmarks
- SSSP for each Landmark
- Ball $B(u)$ for each Node u (expected size $\mathcal{O}(\sqrt{n})$)

$$B(u) := \{v \in V(G) : d_G(u, v) < d_G(u, \ell_u)\},$$

  where ℓ_u denotes u's nearest landmark

Query $d(u, v)$

- If $v \in B(u)$ (or vice versa $u \in B(v)$) return exact distance
- Otherwise $u \to \ell_v \to v$, where ℓ_v denotes v's nearest landmark

Modified Thorup Zwick

Preprocessing

- $\mathcal{O}(n^\gamma)$ High-degree Nodes as Landmarks
- SSSP for each Landmark
- Ball $B(u)$ for each Node u (expected size $\mathcal{O}(n^\gamma)$)

$$B(u) := \{v \in V(G) : d_G(u, v) < d_G(u, \ell_u)\},$$

where ℓ_u denotes u's nearest landmark

Query $d(u, v)$

- If $v \in B(u)$ (or vice versa $u \in B(v)$) return exact distance
- Otherwise $u \to \ell_v \to v$, where ℓ_v denotes v's nearest landmark

Idea of Proof that Balls have Size at most $\mathcal{O}(n^\gamma)$

Idea: when growing a ball from u,...

- either detect a node with high degree $\rightsquigarrow$ ✓
  - Chung-Lu model: node degree is proportional to a potential value
  - Nodes are connected based on this potential value
- or all nodes have small degree $\rightsquigarrow$ small ball growth ✓

Distance Oracles for Power-law Graphs

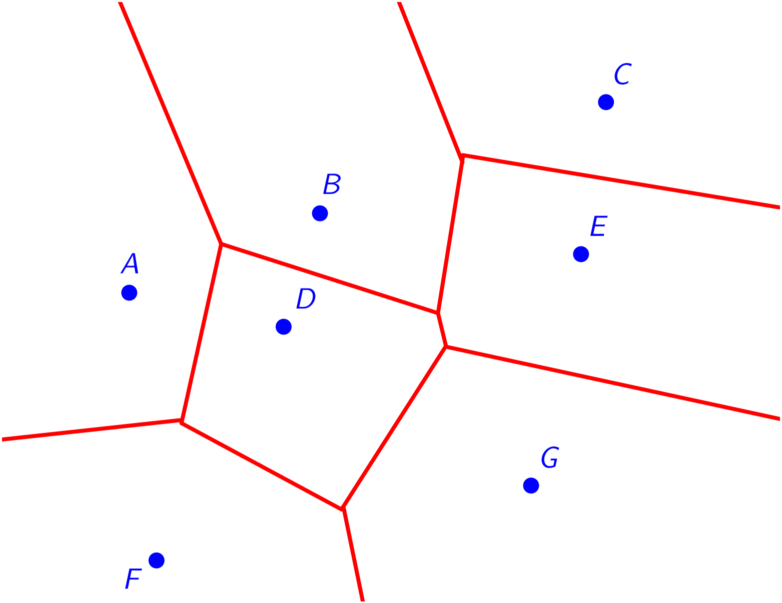
- High-Degree Heuristic is Provably Good for Random Power-law Graphs
- First Approach without Separators (or similar)
- Efficient Preprocessing and Low Space Consumption (significantly better than [TZ05] for general graphs)
- Multiplicative (worst-case) Stretch 3

Outline

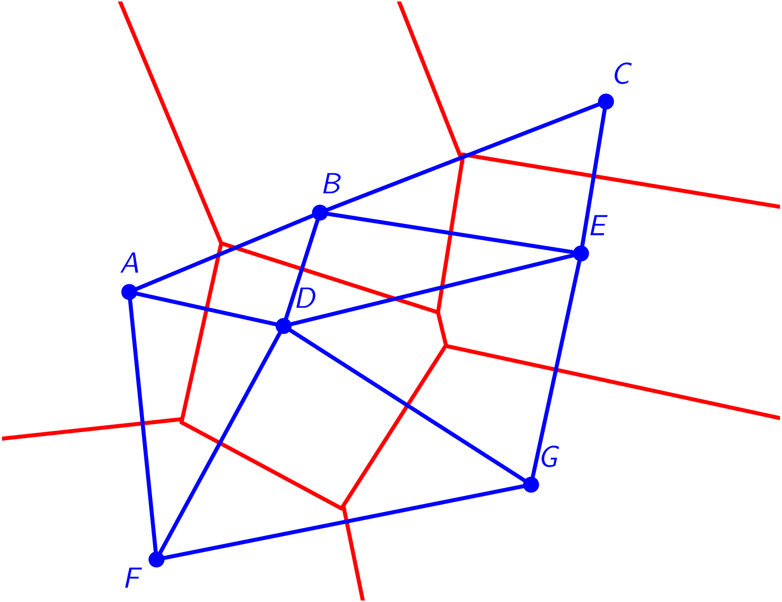
A Practical Method

- Works for a variety of graphs
  - Related work: mostly transportation networks
  - Previous result: high-degree heuristic fails for general networks, still $\mathcal{O}(n^{1+\epsilon})$ preprocessing and space
- Efficient
  - fast preprocessing
  - significantly faster than Dijkstra's algorithm to answer queries
- Simple

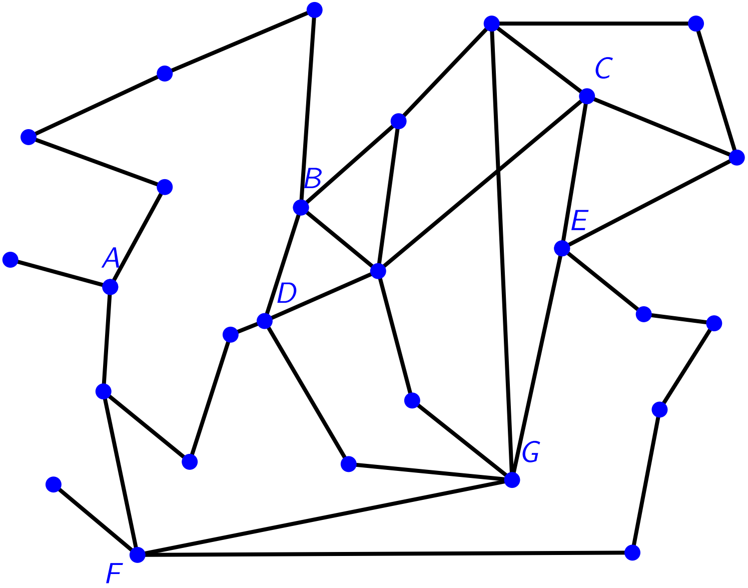
Voronoi Diagram



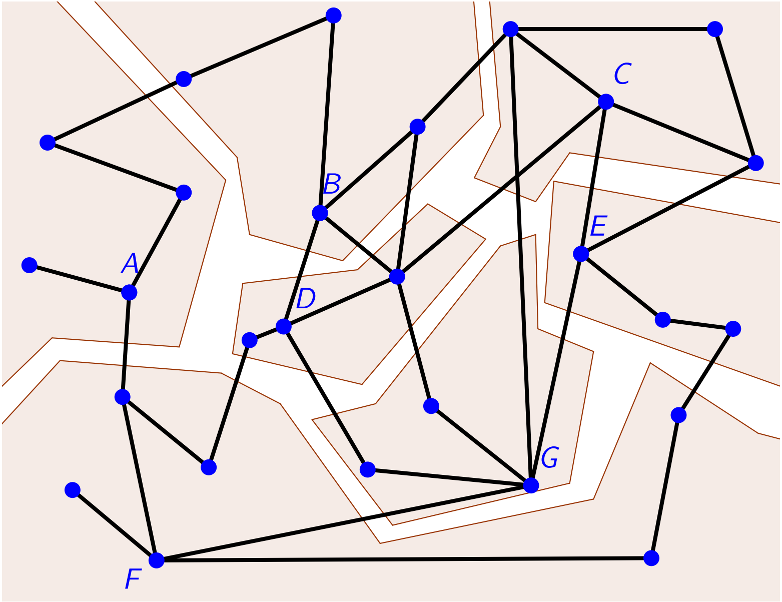
Voronoi Dual: Delaunay Triangulation



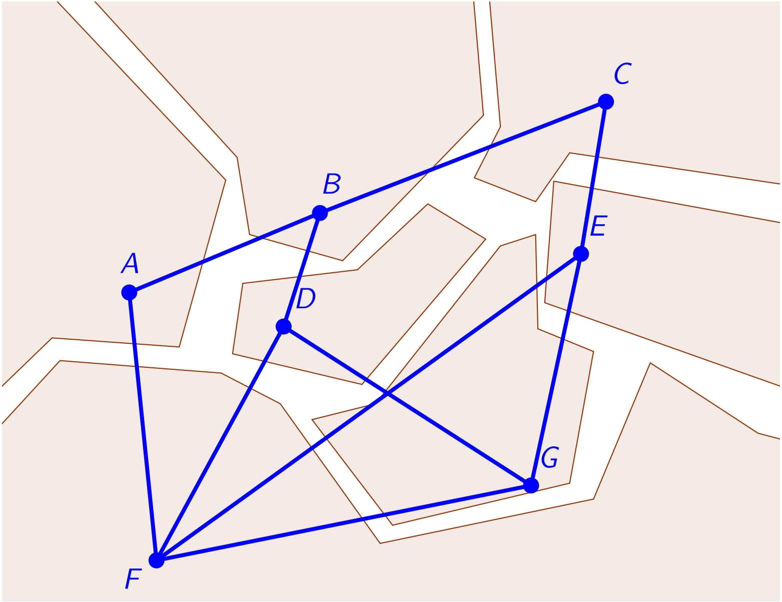
Voronoi Diagram for a Graph?



Graph Voronoi Diagram



Graph Voronoi Dual — not necessarily triangulated



Graph Voronoi Diagram

- Mehlhorn [Meh88] and Erwig [Erw00]
- Can be constructed in time $\mathcal{O}(m + n \lg n)$ using Dijkstra's algorithm once $\leadsto$ very efficient

- Graph Voronoi Dual can also be constructed using one SSSP search

Voronoi Approximate Shortest Path Query Method

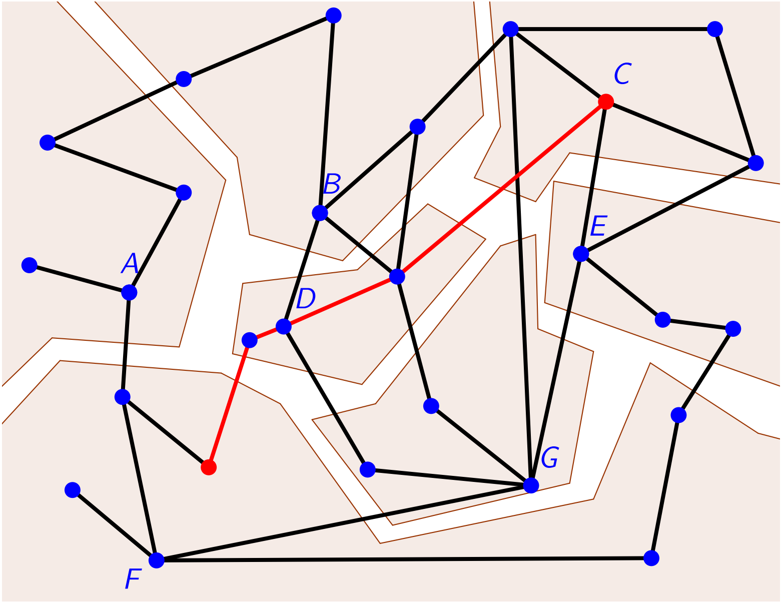
Preprocessing

- Random Sampling of Voronoi Nodes
- Voronoi Dual

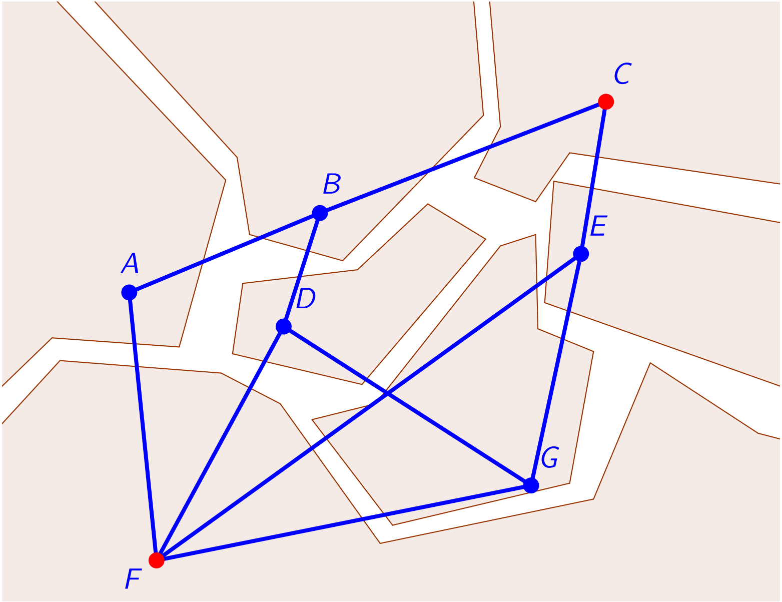
Query $d(u, v)$

- Shortest Path between $\text{vor}(u)$ and $\text{vor}(v)$ in Voronoi Dual
- (Optional) Refinement: Shortest Path in Voronoi Regions

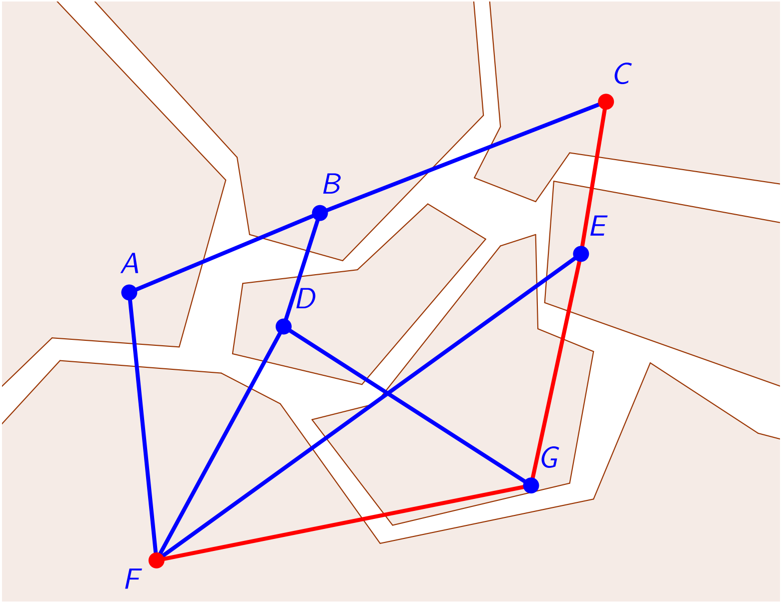
Voronoi Method — Query



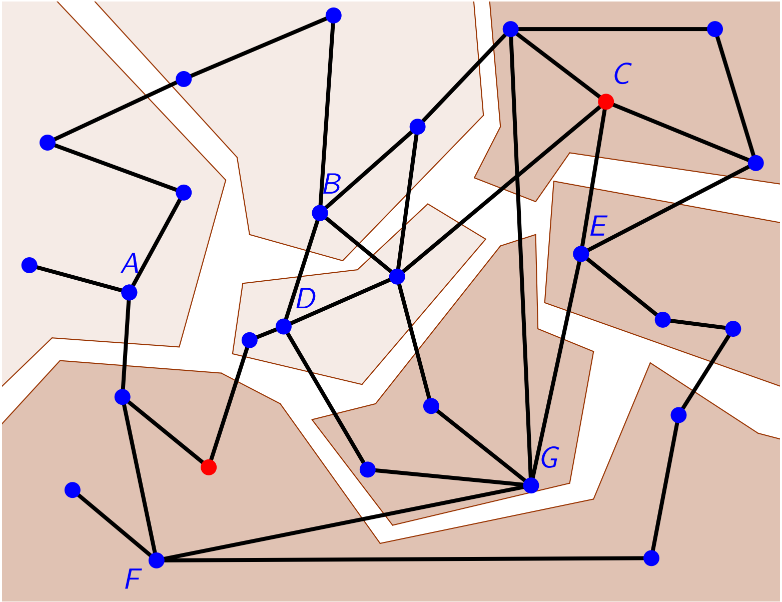
Voronoi Method — Query: Search in Dual



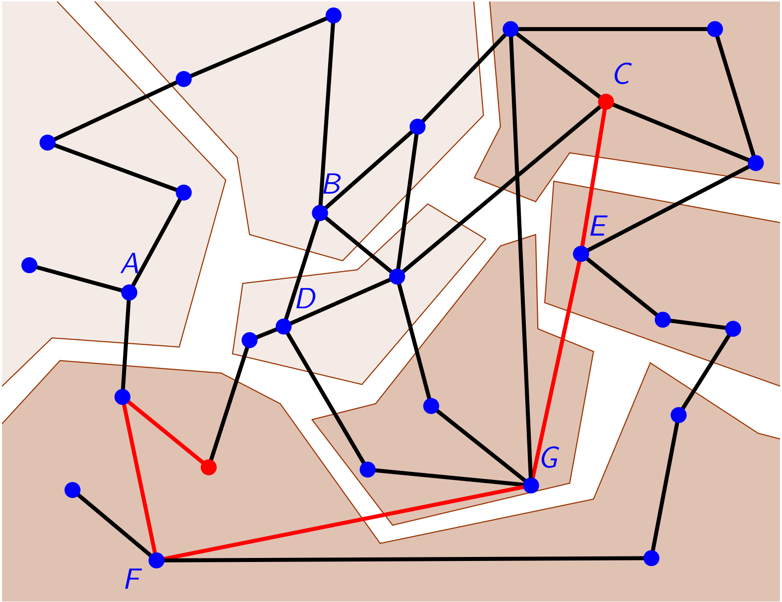
Voronoi Method — Query: Search in Dual



Voronoi Method — Query: Search in ⊆ Primal



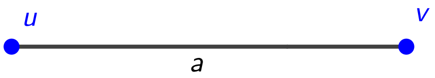
Voronoi Method — Query Result



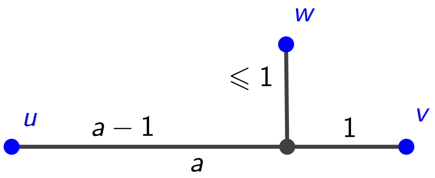
Graph Voronoi Diagram — Stretch?

- Theoretical bound: expected $\mathcal{O}\left(\frac{\lg \ell}{\lg \frac{1}{1-p}}\right)$
  - for paths with ℓ edges
  - sampling probability p
- Experimental
  - $1+\epsilon$ for road networks
  - less than 2 for complex networks
  - does not depend on ℓ

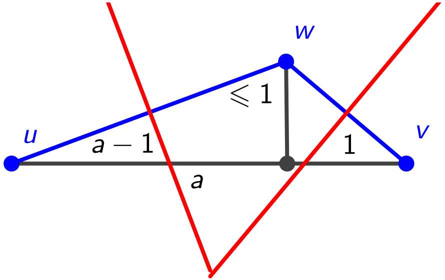
Voronoi Method — Stretch



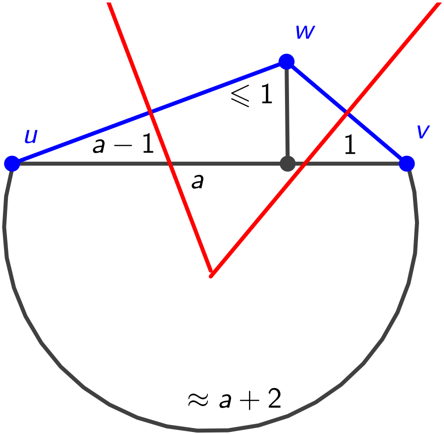
Voronoi Method — Stretch



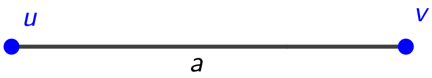
Voronoi Method — Stretch



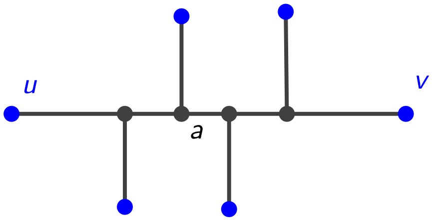
Voronoi Method — Stretch



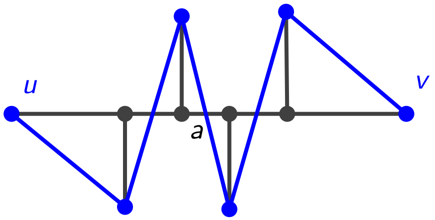
Voronoi Method — Stretch



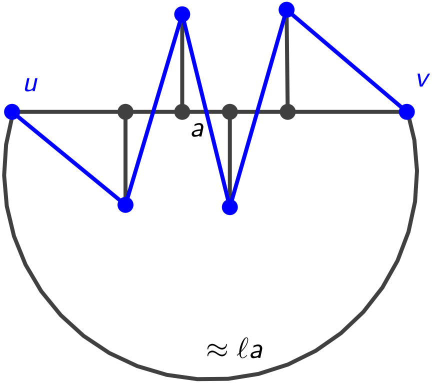
Voronoi Method — Stretch



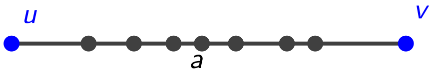
Voronoi Method — Stretch



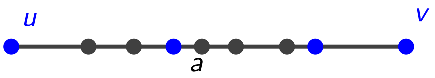
Voronoi Method — Stretch



Voronoi Method — Stretch



Voronoi Method — Stretch



expected $\mathcal{O}\left(\frac{\lg \ell}{\lg \frac{1}{1-p}}\right)$

Summary

- Very efficient preprocessing
- Competitive query time
- Simple
- ... but not exact
- Can trade query time vs. stretch

- Average-case stretch?

Outline

- Introduction
- Space Lower Bound
- Distances in Power-Law Graphs
- A Practical Method
- Conclusion

Conclusion

Contribution of this thesis

- Proof that there is no efficient ($o(\lg n)$ query time) distance oracle with space $\mathcal{O}(m)$
- [TZ05] works provably well for power-law graphs
- There is a practically efficient distance oracle with space $\mathcal{O}(m)$

Thank you!

Thank you!

Rey Abe, Ittai Abraham, Hikari Aikawa, Alicia Aponte, Yuki Arase, Cyrille Artho, Jörg Aschwanden, David M. Avis, Yukino Baba, Valentina Baljak, Nik Nailah Binti Abdullah, Guillaume Bouvet, Samuel Burri, Ulrike Campbell & Glenn, Nicholas Campiz, Gary & Mami Chandler, Kun Chen, Wei Chen, Kazuo Chiba, Chad Clark, Edward Clease, Leticia Cuellar, Erik D. Demaine, Joachim Dirks, Hristo Djidjev, Takuo Doi, Adrian Doswald, Sébastien Duval, Stephan Eidenbenz (and family), Jittat Fakcharoenphol, Daisuke Fukuchi, Pierre-Loïc Garoche, Cyril Gavoille, Filiz Gencer, Jacopo Grazzini, Nizar Grira, Leonid Gurvits, Daniel & Monika Hagemeier, Martin Halter, Rie Hatanaka (and family), Nicolas W. Hengartner, Roger Herzog, Katsushige Hino, Jon Holmes, Shinichi Honiden, Michael E. Houle, Jaimie Hwang, Takeo Igarashi, Hiroshi Imai, Julia Imhof, Mizuki Inoue, Rihoko Inoue, Fuyuki Ishikawa, Katsuro Ishimasa, Tanja Isker, Moritz Isler, Nicolas Jegou, Jay Johannesen, Chika Kanai, Shiva Kasiviswanathan, Satoshi Kataoka, Douglas D. Kautz, Yojiro Kawamata, Ken-ichi Kawarabayashi, Daniel Kershner, Junichi 'Andy' Kimura, Jumi Klaus, Adrian Klein, Gordon Kanki Knight, Yusuke Kobayashi, Lukas Kroc, Mitsuhiko & Yuko Kusuyama, Stephen Lacey, Christiane Lange, Stephan Langermann, Brett Larner, Jason Lawrence, Xiang-Yang Li, Maxim Makatchev, Carrie Manore, Koichi Matsumoto, Yasuyuki Matsushita, Kazutaka Matsuzaki, Petar Maymounkov, John & Yuki Mettraux, Martin Mevissen, Omar Minami, Teruyuki Minegishi, Keren Miers, Felix Mösner, Hirotaka Moriguchi, Mohammad Reza Motallebi, Machiko & Yoshinobu Nagura, Hiroyuki & Miki Nakagawa (and Family Nishida in Takatsuki), Yoshiyuki Nakamura, Hikotoshi Nakazato, Michael Nett, Satohi Numasawa, Johan Nyström, Paddy O'Connor, Kyoko Oda, Akemi Okado, Chieko Okamoto, Yoshio Okamoto, Masae Ono, Rie Onodera, Takeshi Ozawa, Eric Platon, Bob Poulson, Hans Prisi, Gareth Pughe, José Ghislain Quenum, Daniele Quercia, Fabrizio Raponi, Jerry Ray, Martin Rehak, Gerard Robb, David Roberts, David Motozo Rubenstein, Cedric S. Rutishauser, Philip Ryan, Kunihiko Sadakane, Hitomi Sakamoto, Nandakishore Santhi, Chikako, Tomoaki, & Noriaki Sawada, Tomoko Sawada, Yuzuru Sawato, Christoph Saxer, Hannes Schneebeli, Gabriel Schweizer, Yuichi Sei, Ryota Seike, Tetsuo Shibuya, Akiko Shimazu, Shunichiro Suenaga, Yuka Shigihara, Uwe Sievers, Weihuan Shu, Lolan Song, Hansruedi, Hermine, and Stefan Sommer, Thomas, Kati & Lars Spirig, Toshiro 'Joe' Suzuki, Kenji Taguchi, Yasuyuki Tahara, Ryuichi Takahashi, Keren Tan, Ryu Tatsumi, 'Tommy', Kenji Tei, Shang-Hua Teng, Mikkel Thorup, Sunil Thulasidasan, Ai Tobimatsu, Mika Tokairin, Shigeko Tokuda, Susumu Toriumi, Georg Troxler, Eric Tschetter, Ryudo Tsukizaki, Takeaki Uno, Elad Verbin, Martin Verdier, Vishwanath Venkatesan, Yajun Wang, Martin Wenk, Hermann Werner, Stefan Wolf, Martin Wolff, Jürgen Wittstock, Shuko Yamada, Ken Yamagata, Kayoko Yamamoto, Guanhua Yan, Andrew Yao, Kiyonari Yoshida, Nobukazu Yoshioka, Wei Yu, Jialin Zhang, Yuan Zhou, Peter & Ursi Zürcher, Uri Zwick

Ittai Abraham and Cyril Gavoille.
Object location using path separators.
In Proceedings of the Twenty-Fifth Annual ACM Symposium on Principles of Distributed Computing, PODC 2006, Denver, CO, USA, July 23-26, 2006, pages 188–197, 2006.
Details in LaBRI Research Report RR-1394-06.

Reinhard Bauer and Daniel Delling.
SHARC: Fast and robust unidirectional routing.
In Proceedings of the 10th Workshop on Algorithm Engineering and Experiments (ALENEX'08), pages 13–26, 2008.

Reinhard Bauer, Daniel Delling, Peter Sanders, Dennis Schieferdecker, Dominik Schultes, and Dorothea Wagner.
Combining hierarchical and goal-directed speed-up techniques for Dijkstra's algorithm.
In Experimental Algorithms, 7th International Workshop (WEA'08), Provincetown, MA, USA, May 30-June 1, 2008, Proceedings, pages 303–318, 2008.

Holger Bast, Stefan Funke, Peter Sanders, and Dominik Schultes.
Fast routing in road networks with transit nodes.
Science, 316(5824):566, 2007.

Surender Baswana, Akshay Gaur, Sandeep Sen, and Jayant Upadhyay.
Distance oracles for unweighted graphs: Breaking the quadratic barrier with constant additive error.
In Automata, Languages and Programming, 35th International Colloquium, ICALP 2008, Reykjavik, Iceland, July 7-11, 2008, Proceedings, Part I: Tack A: Algorithms, Automata, Complexity, and Games, pages 609–621, 2008.

Surender Baswana and Telikepalli Kavitha.
Faster algorithms for approximate distance oracles and all-pairs small stretch paths.
In 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006), 21-24 October 2006, Berkeley, California, USA, pages 591–602, 2006.

Surender Baswana and Sandeep Sen.
Approximate distance oracles for unweighted graphs in expected $\mathcal{O}(n^2)$ time.
ACM Transactions on Algorithms, 2(4):557–577, 2006.
Announced at SODA 2004.

Shiva Chaudhuri and Christos D. Zaroliagis.
Shortest paths in digraphs of small treewidth. part I: Sequential algorithms.

Algorithmica, 27(3):212–226, 2000.
Announced at ICALP 1995.

Martin Erwig.
The graph Voronoi diagram with applications.
Networks, 36(3):156–163, 2000.

Andrew V. Goldberg and Chris Harrelson.
Computing the shortest path: A* search meets graph theory.
In Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'05), Vancouver, British Columbia, Canada, January 23-25, 2005, pages 156–165, 2005.

Dmitri V. Krioukov, Kevin R. Fall, and Xiaowei Yang.
Compact routing on internet-like graphs.
In INFOCOM, 2004.

Philip Nathan Klein.
Preprocessing an undirected planar network to enable fast approximate distance queries.
In Symposium on Discrete Algorithms (SODA), pages 820–827, 2002.

Alexander Lubotzky, R. Phillips, and Peter Sarnak.
Ramanujan graphs.
Combinatorica, 8(3):261–277, 1988.

Kurt Mehlhorn.
A faster approximation algorithm for the Steiner problem in graphs.
Information Processing Letters, 27(3):125–128, 1988.

Manor Mendel and Assaf Naor.
Ramsey partitions and proximity data structures.
In 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006), 21-24 October 2006, Berkeley, California, USA, Proceedings, pages 109–118, 2006.

Manor Mendel and Chaya Schwob.
C-K-R partitions of sparse graphs.
CoRR, abs/0809.1902, 2008.

Mihai Patrascu.
(Data) STRUCTURES.
In 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA, pages 434–443, 2008.

Mihai Patrascu and Mikkel Thorup.
Time-space trade-offs for predecessor search.
In Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006, pages 232–240, 2006.

Liam Roditty, Mikkel Thorup, and Uri Zwick.
Deterministic constructions of approximate distance oracles and spanners.
In Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Lisbon, Portugal, July 11-15, 2005, Proceedings, pages 261–272, 2005.

Peter Sanders and Dominik Schultes.
Highway hierarchies hasten exact shortest path queries.
In Algorithms - ESA 2005, 13th Annual European Symposium, Palma de Mallorca, Spain, October 3-6, 2005, Proceedings, pages 568–579, 2005.

Robert Sedgewick and Jeffrey Scott Vitter.
Shortest paths in euclidean graphs.
Algorithmica, 1(1):31–48, 1986.
Announced at FOCS 1984.

Kunal Talwar.
Bypassing the embedding: algorithms for low dimensional metrics.
In STOC '04: Proceedings of the thirty-sixth annual ACM symposium on Theory of computing, pages 281–290, 2004.

Mikkel Thorup.
Compact oracles for reachability and approximate distances in planar digraphs.

Journal of the ACM, 51(6):993–1024, 2004.
Announced at FOCS 2001.

Mikkel Thorup and Uri Zwick.
Approximate distance oracles.
Journal of the ACM, 52(1):1–24, 2005.
Announced at STOC 2001.