Load Balanced Routing Using Area Preserving Maps

Jie Gao

Stony Brook University

Dagstuhl Seminar, Jan 27-31, 2014.

Load Balanced Routing

Given source destination pairs, find routes that minimize the maximum traffic load any node carries.

Load Balanced Routing

Given source destination pairs, find routes that minimize the maximum traffic load any node carries.

- Congestion causes long delay

Load Balanced Routing

Given source destination pairs, find routes that minimize the maximum traffic load any node carries.

- Congestion causes long delay
- Nodes running out of battery.

Load Balanced Routing

Given source destination pairs, find routes that minimize the maximum traffic load any node carries.

- Congestion causes long delay
- Nodes running out of battery.

Maximum traffic load depends on

Load Balanced Routing

Given source destination pairs, find routes that minimize the maximum traffic load any node carries.

- Congestion causes long delay
- Nodes running out of battery.

Maximum traffic load depends on

- Traffic pattern: distribution of source and destination pairs. E.g., all-to-one or one-to-all.

Load Balanced Routing

Given source destination pairs, find routes that minimize the maximum traffic load any node carries.

- Congestion causes long delay
- Nodes running out of battery.

Maximum traffic load depends on

- Traffic pattern: distribution of source and destination pairs. E.g., all-to-one or one-to-all.
- Network topology, e.g., ‘narrow neck’.

Load Balanced Routing

Given source destination pairs, find routes that minimize the maximum traffic load any node carries.

- Congestion causes long delay
- Nodes running out of battery.

Maximum traffic load depends on

- Traffic pattern: distribution of source and destination pairs. E.g., all-to-one or one-to-all.
- Network topology, e.g., 'narrow neck'.
- Imperfect routing algorithm.

Load Balanced Routing on Graphs

1. The unsplittable flow problem: select routes that minimize max load.
   - NP-hard to approximate within $\Omega(\log^{1/2-\varepsilon} n)$. [Andrews, Zhang 07]
   - Approximation algorithm w. factor $O(\log n / \log\log n)$. [Raghavan 88]

Load Balanced Routing on Graphs

1. The unsplittable flow problem: select routes that minimize max load.
   - NP-hard to approximate within $\Omega(\log^{1/2-\varepsilon} n)$. [Andrews, Zhang 07]
   - Approximation algorithm w. factor $O(\log n / \log\log n)$. [Raghavan 88]
2. Find node or edge disjoint paths that deliver max # source destination pairs.
   - NP-hard to approximate within $\Omega(\log^{1/2-\varepsilon} n)$. [Chuzhoy, Khanna 05]
   - Approximation algorithm w. $O(\sqrt{n})$. [Chekuri, Khanna, Shepherd 06]

Load Balanced Routing on Graphs

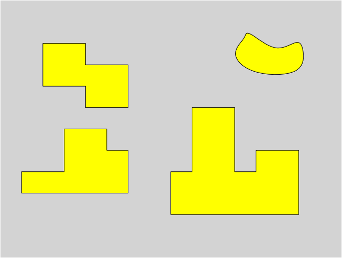
1. The unsplittable flow problem: select routes that minimize max load.
   - NP-hard to approximate within $\Omega(\log^{1/2-\varepsilon} n)$. [Andrews, Zhang 07]
   - Approximation algorithm w. factor $O(\log n / \log\log n)$. [Raghavan 88]
2. Find node or edge disjoint paths that deliver max # source destination pairs.
   - NP-hard to approximate within $\Omega(\log^{1/2-\varepsilon} n)$. [Chuzhoy, Khanna 05]
   - Approximation algorithm w. $O(\sqrt{n})$. [Chekuri, Khanna, Shepherd 06]

Our setting: a dense set of nodes deployed inside a geometric domain.

Modeling the Geometry of Wireless Sensor Networks

Large-scale dense deployment of n sensors in a domain Ω.

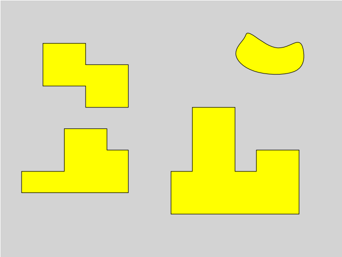
- Modeling the large scale domain features (obstacles, lake, etc), whose size $\gg$ communication range.



Modeling the Geometry of Wireless Sensor Networks

Large-scale dense deployment of n sensors in a domain Ω.

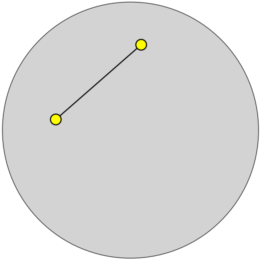
- Modeling the large scale domain features (obstacles, lake, etc), whose size $\gg$ communication range.



Goal: how does the geometric shape influence load balancing, with given traffic pattern (joint distribution Π of source destination)?

Examples

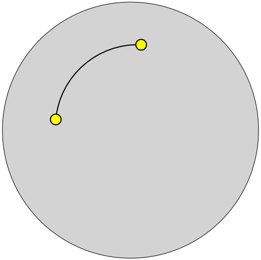
Shortest path routing in a disk shape network with uniform traffic.



Traffic load at the center is the highest.

Examples

Shortest path routing in a disk shape network with uniform traffic.



Push routes to the boundary. Path stretch $\uparrow$. Traffic load at the center $\downarrow$.

Trade-off Between Path Stretch and Load Balancing

Path stretch α: $\frac{\text{path length}}{\text{shortest path length}}$.

Trade-off Between Path Stretch and Load Balancing

Path stretch α: $\frac{\text{path length}}{\text{shortest path length}}$.

Load balancing ratio β: $\frac{\text{max load}}{\text{max load in OPT}}$.

Trade-off Between Path Stretch and Load Balancing

Path stretch α: $\frac{\text{path length}}{\text{shortest path length}}$.

Load balancing ratio β: $\frac{\text{max load}}{\text{max load in OPT}}$.

Q: Can we get $O(1)$ path stretch and $O(1)$ load balancing ratio?

Trade-off Between Path Stretch and Load Balancing

Path stretch α: $\frac{\text{path length}}{\text{shortest path length}}$.

Load balancing ratio β: $\frac{\text{max load}}{\text{max load in OPT}}$.

Q: Can we get $O(1)$ path stretch and $O(1)$ load balancing ratio?

- NO in general: $\beta = \Theta(\sqrt{n/\alpha})$ in constant-degree unit disk graphs. [Gao, Zhang 04]

Trade-off Between Path Stretch and Load Balancing

Path stretch α: $\frac{\text{path length}}{\text{shortest path length}}$.

Load balancing ratio β: $\frac{\text{max load}}{\text{max load in OPT}}$.

Q: Can we get $O(1)$ path stretch and $O(1)$ load balancing ratio?

- NO in general: $\beta = \Theta(\sqrt{n/\alpha})$ in constant-degree unit disk graphs. [Gao, Zhang 04]
- YES if the network has a 'nice' shape.

Scaling Law for Shortest Paths in Sphere Network

- n nodes uniformly on a **sphere**. $O(1)$ degree.
- One message for each pair of nodes.

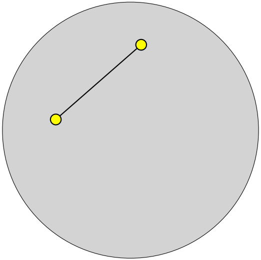
Scaling Law for Shortest Paths in Sphere Network

- n nodes uniformly on a **sphere**. $O(1)$ degree.
- One message for each pair of nodes.

Avg path length $\Theta(\sqrt{n})$. Total traffic $O(n^2\sqrt{n})$. $\Rightarrow$ Traffic load $\Theta(n\sqrt{n})$ everywhere.

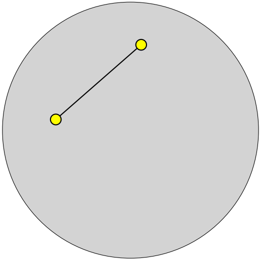
Scaling Law for Shortest Paths in Disk Network

Maximum traffic load $= \Theta(n\sqrt{n})$ [Jonckheere, Lou, Bonahon, Baryshnikov 11]



Scaling Law for Shortest Paths in Disk Network

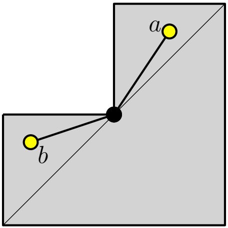
Maximum traffic load = $\Theta(n\sqrt{n})$ [Jonckheere, Lou, Bonahon, Baryshnikov 11]



$\alpha = 1$, $\beta = O(1) \Rightarrow$ asymptotically optimal.

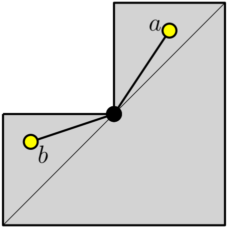
Shortest Paths in the L shape

The reflex vertex has traffic load $\Theta(n^2)$.



Shortest Paths in the L shape

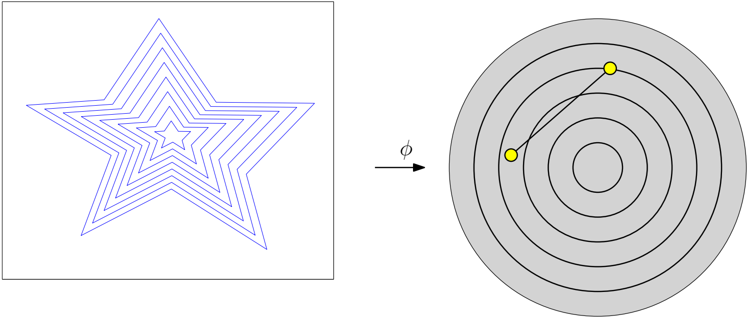
The reflex vertex has traffic load $\Theta(n^2)$.



$\alpha = 1$, $\beta = \Theta(\sqrt{n}) \Rightarrow$ the worst suggested by the trade-off analysis.

Our Approach

- Area Preserving Map $\phi : \Omega \to D$ disk.
- Use virtual coordinates on the disk for greedy routing.

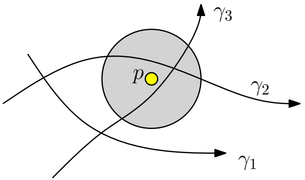


Definition of Traffic Load in Continuous Setting

Def: Traffic load $\ell(A)$ of any region $A \subseteq \Omega$.

- Choose k source destination pairs (a_i, b_i) from the traffic distribution Π.
- Route γ_i connecting a_i to b_i.
- $X(A)$: average length of all paths inside A.

$$\ell(A) = \lim_{k \to \infty} \frac{X(A)}{Area(A)}$$

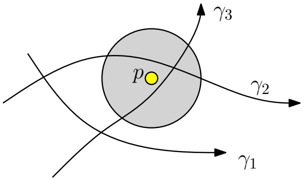


Definition of Traffic Load in Continuous Setting

Def: Traffic load $\ell(p)$ at **point** $p \in \Omega$.

- Choose k source destination pairs (a_i, b_i) from the traffic distribution Π.
- Route γ_i connecting a_i to b_i.
- **Choose a nested series of neighborhood A_j including p, $\lim_{j\to\infty} Area(A_j) \to 0$.**
- $X(A_j)$: average length of all paths inside A_j.

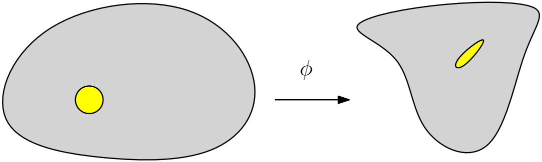
$$\ell(p) = \lim_{j\to\infty} \ell(A_j)$$



Area Preserving Map

Given Ω, Ω' in $\mathbb{R}^n$, $\phi : \Omega \to \Omega'$ is area-preserving iff for any $A \subseteq \Omega$,

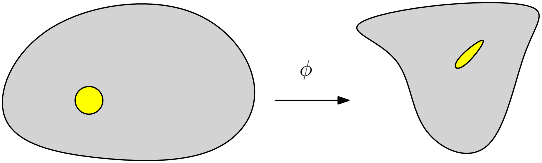
$$Area(A) = Area(\phi(A))$$



Area Preserving Map

Given Ω, Ω' in $\mathbb{R}^n$, $\phi : \Omega \to \Omega'$ is area-preserving iff for any $A \subseteq \Omega$,

$$Area(A) = Area(\phi(A))$$

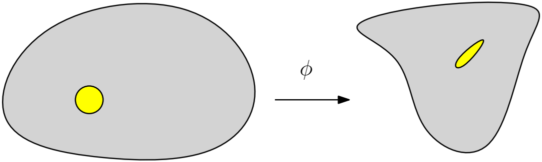


- Jacobian matrix J has determinant 1 everywhere.

Area Preserving Map

Given Ω, Ω' in $\mathbb{R}^n$, $\phi : \Omega \to \Omega'$ is area-preserving iff for any $A \subseteq \Omega$,

$$Area(A) = Area(\phi(A))$$

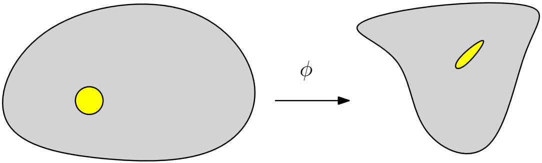


- Jacobian matrix J has determinant 1 everywhere.
- Eigenvalues $\lambda_1(p), \lambda_2(p)$: $\lambda_1(p) \cdot \lambda_2(p) = 1$.

Area Preserving Map

Given Ω, Ω' in $\mathbb{R}^n$, $\phi : \Omega \to \Omega'$ is area-preserving iff for any $A \subseteq \Omega$,

$$Area(A) = Area(\phi(A))$$

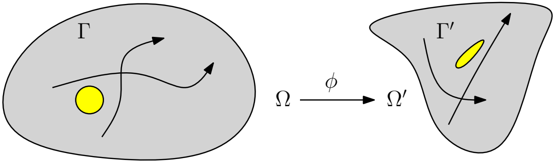


- Jacobian matrix J has determinant 1 everywhere.
- Eigenvalues $\lambda_1(p), \lambda_2(p)$: $\lambda_1(p) \cdot \lambda_2(p) = 1$.
- Maximum length distortion of ϕ: $d = \max_p(\lambda_1(p), \lambda_2(p))$

Traffic Load Under ϕ

Theorem: Given a routing algorithm Γ on Ω and $\phi : \Omega \to \Omega'$ an area preserving map, the traffic load at p by routing scheme $\Gamma' = \phi(\Gamma)$ is

$$\frac{1}{d(p)}\ell(p) \leq \ell'(\phi(p)) \leq d(p)\ell(p)$$

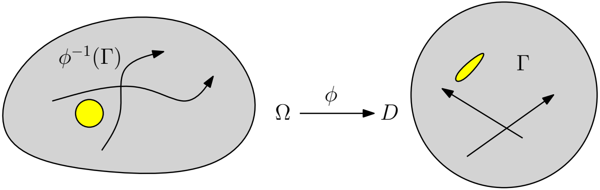


Proof: Recall that $\ell(A) = \lim_{k\to\infty} \frac{X(A)}{Area(A)}$.

Load Balanced Short Path Routing in Ω

Theorem: Given an area preserving map $\phi : \Omega \to D$ with maximum distortion d, and a routing scheme Γ on D with $O(1)$ stretch and $O(1)$ load balancing ratio, its pullback $\phi^{-1}(\Gamma)$ has:

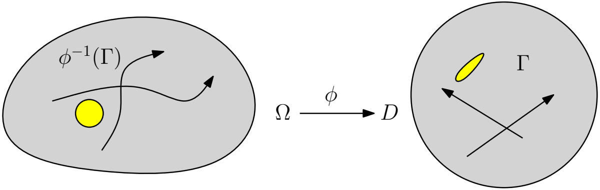
- Path stretch $O(d^2)$.
- Load balancing ratio $O(d^2)$.



Load Balanced Short Path Routing in Ω

Theorem: Given an area preserving map $\phi : \Omega \to D$ with maximum distortion d, and a routing scheme Γ on D with $O(1)$ stretch and $O(1)$ load balancing ratio, its pullback $\phi^{-1}(\Gamma)$ has:

- Path stretch $O(d^2)$.
- Load balancing ratio $O(d^2)$.



Why d^2? Our path can be stretched d times longer; while optimal path get stretched d times shorter.

What Next?

- Find an area preserving map ϕ from Ω to D.

What Next?

- Find an area preserving map ϕ from Ω to D.
- Find load balanced short path routing in the disk.

Improving Over Shortest Path on Disk

Map the disk to sphere. $\uparrow$ path stretch s.t. max load $\downarrow$.

Improving Over Shortest Path on Disk

Map the disk to sphere. $\uparrow$ path stretch s.t. max load $\downarrow$.

- Stereographic projection: angle-preserving. [Curveball routing, Popa, Rostamizadeh, Karp, Papadimitriou, Stoica 07]

Improving Over Shortest Path on Disk

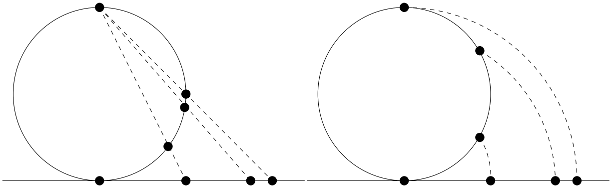
Map the disk to sphere. $\uparrow$ path stretch s.t. max load $\downarrow$.

- Stereographic projection: angle-preserving. [Curveball routing, Popa, Rostamizadeh, Karp, Papadimitriou, Stoica 07]
- Lambert Azimuthal projection: area-preserving.

Improving Over Shortest Path on Disk

Map the disk to sphere. ↑ path stretch s.t. max load ↓.

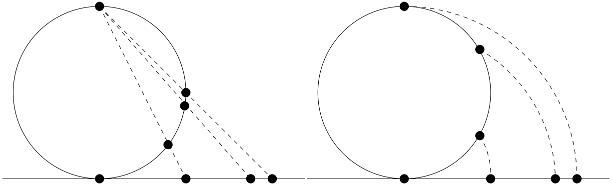
- Stereographic projection: angle-preserving. [Curveball routing, Popa, Rostamizadeh, Karp, Papadimitriou, Stoica 07]
- Lambert Azimuthal projection: area-preserving.



Improving Over Shortest Path on Disk

Map the disk to sphere. $\uparrow$ path stretch s.t. max load $\downarrow$.

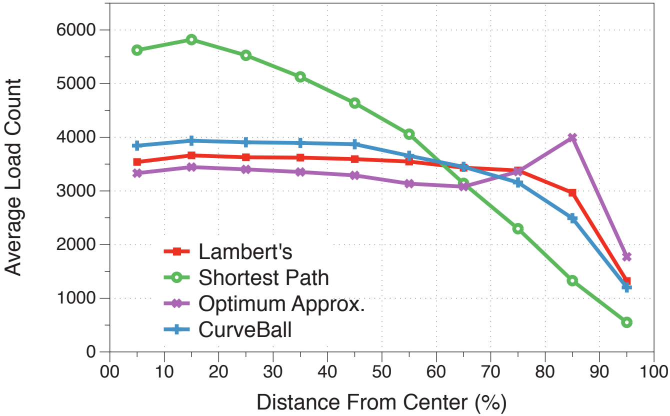
- Stereographic projection: angle-preserving. [Curveball routing, Popa, Rostamizadeh, Karp, Papadimitriou, Stoica 07]
- Lambert Azimuthal projection: area-preserving.



Theorem: Shortest path using virtual spherical coordinate through Lambert Azimuthal projection has stretch of 2 and maximum load $\frac{1}{4\sqrt{2}}$ that of shortest path routing in the disk.

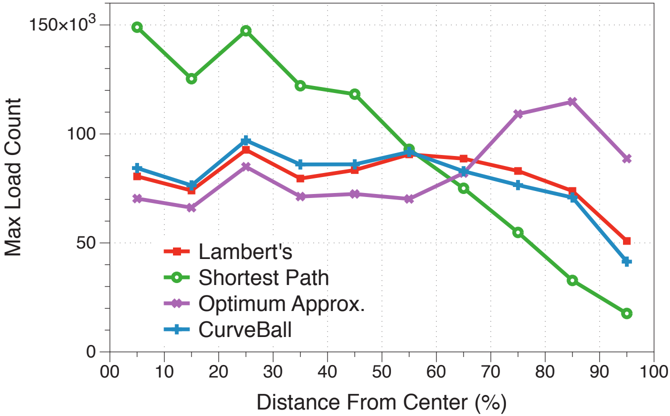
Routing in a Disk

Avg load for different routing schemes in a disk.



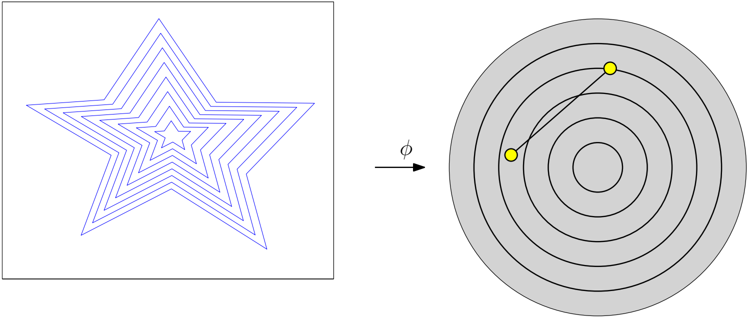
Routing in a Disk

Max load for different routing schemes in a disk.



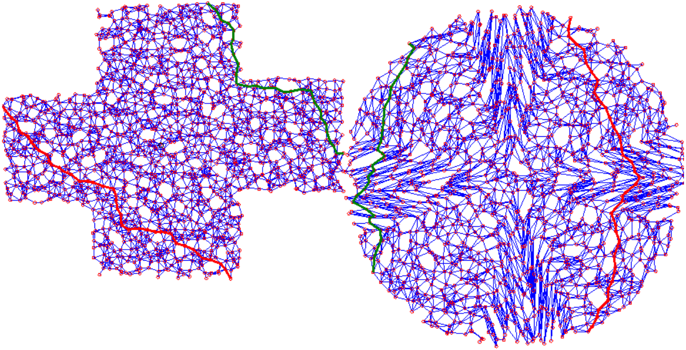
Find an Area Preserving Map

- Find a contour generating function f in Ω.
- Map contours to concentric circles. [Brown, Halperin 35]



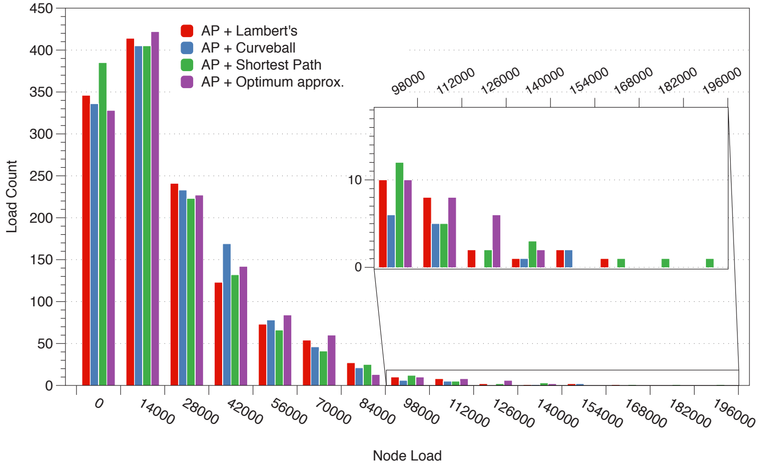
Routing in a Cross Shape

Area preserving map



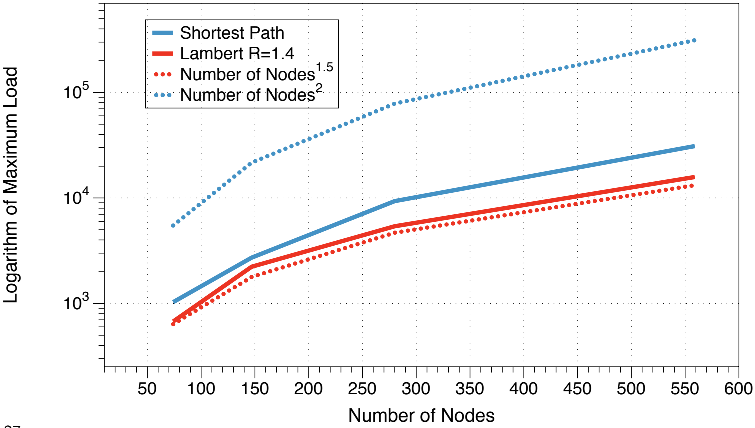
Routing in a Cross Shape

Histogram of traffic load, using Lambert projection to sphere.



Scaling of Max Load in L-shape

Max load in original coordinates (scales $\sim n^2$) and by our method (scales $\sim n^{1.5}$).



Conclusion and Ongoing Work

Shape of the network matters in load balancing!

Conclusion and Ongoing Work

Shape of the network matters in load balancing!

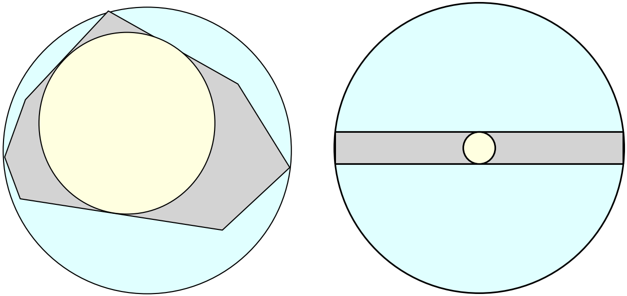
How to characterize domain Ω that admits an area preserving map to D with small distortion?

Conclusion and Ongoing Work

Shape of the network matters in load balancing!

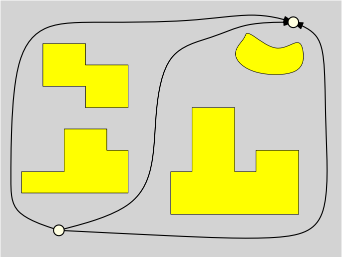
How to characterize domain Ω that admits an area preserving map to D with small distortion?

- Convex domain: fatness, $\dfrac{\text{Radius of maximum inscribing disk}}{\text{Radius of minimum enclosing disk}}$



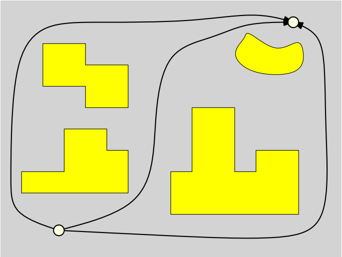
Conclusion and Ongoing Work

Non-simple domain?



Conclusion and Ongoing Work

Non-simple domain?



Load Balancing at two levels

- Top-level: decide on homotopy types.
- Bottom-level: spread out traffic.

Acknowledgement

- Joint work with Mayank Goswami, Chien-Chun Ni, Xiaomeng Ban, Xianfeng David Gu, Stony Brook University.
- Questions and comments?