Spring, 2011

Class 16
Routing

• Choose the path on which a message from u to v will be sent

• The goal is to estimate the latency between each pair of nodes

• Consider a complete graph
  – But edges have different latencies
Today

• Triangulation and embedding using small sets of beacons, Kleinberg, Slivkins and Wexler

• Towards fast decentralized construction of locality-aware overlay networks, Slivkins
Motivation

• Calculating latencies (round-trip) in a Network
• Networking algorithms calculate latency
  – Cannot measure all distances
• Latency is almost a Metric
  – Non-negative: \( d(x,y) \geq 0 \) and \( d(x,x)=0 \)
  – Symmetric: \( d(x,y) = d(y,x) \)
  – Triangle inequality: \( d(x,y) \leq d(x,z)+d(z,y) \)
Beacon-Based Approach

• Feasible to measure only a small number of distances
• Choose uniformly at random a constant number of nodes (beacons) and measure all distances to them
• Reconstruction by triangulation
  – Bound other distances by triangle inequality
Reconstruction by Triangulation

- Distance $d_{uv}$ between points $u$ and $v$
- $S$ is the set of beacons

$$\max_{b \in S} |d_{ub} - d_{vb}| \leq d_{uv} \leq \min_{b \in S} d_{ub} + d_{vb}$$

- **Algorithm**: Node $u$ measures distance to all beacons. Exchanges data with node $v$ to estimate $d_{uv}$
- Distortion $1+\delta$ : max of ratio between upper and lower bounds
Slack

• Cannot always approximate all distances
  – Example: for same $d_{uv}$, difference between $d_{ub}$ and $d_{vb}$ can be small, but sum can be small or large

• Slack - approximate with distortion $1+\delta$ all but an $\epsilon$-fraction of distances
Goal

• Given a metric $M$, and $\epsilon, \delta > 0$ : is there a constant number (independent of $|M|$) of beacons that suffice to get triangulation with distortion $1 + \delta$ and slack $\epsilon$?

• Consider $M$ such that all distances are 1
  – Lower bound of triangulation is always 0…

• New goal: find natural metrics for which this can be done
s-doubling Metrics

• $B_u(r) = \{v : d_{uv} \leq r\}$

• In an s-doubling metric: can be covered by at most $2^s$ balls of radius $r/2$
Theorem

• If $M$ is s-doubling, there is a constant number of randomly chosen beacons that achieve distortion $1+\delta$ and slack $\varepsilon$ with probability $1-\gamma$
  – Constant: depends only on $\delta$, $\varepsilon$, $\gamma$, $s$
  – Does not depend on the number of points $|M|$
Proof

• Let \( r_u(\varepsilon) \) be the minimal \( r \) such that \(|B_u(r)| \geq \varepsilon n\)

• Fix \( u \), and \( r = r_u(\varepsilon/3) \). At most \( n\varepsilon/3 \) have distance \(< r\)
  
  – We will ignore the estimation for these points (small \( \varepsilon/3 \) fraction)

• Consider a large ball \( B = B_u(2r/\delta) \)

• The probability of having a beacon in \( B_u(r) \), \( \Pr[\exists b \in B_u(r)] \), is almost 1
  
  for a constant number of random beacons

• Take \( v \notin B \) and check how good will its estimation be:
  
  \[ d_{vb} - d_{ub} \leq d_{uv} \leq d_{vb} + d_{ub} \]
Proof

• Consider a large ball $B = B_u(2r/\delta)$
• Take $v \notin B$ and check how good will its estimation be:
  
  \[ d_{vb} - d_{ub} \leq d_{uv} \leq d_{vb} + d_{ub} \]

• $d_{vb} + d_{ub} \leq (d_{uv} + d_{ub}) + d_{ub} \leq d_{uv} + 2d_{ub} \leq d_{uv} + 2r \leq d_{uv}(1 + 2r/d_{uv}) \leq d_{uv}(1 + \delta)$

• $d_{vb} - d_{ub} \geq (d_{uv} - d_{ub}) - d_{ub} \geq d_{uv} - 2d_{ub} \geq d_{uv} - 2r \geq d_{uv}(1 - 2r/d_{uv}) \geq d_{uv}(1 - \delta)$
Proof

• Now take \( v \in B \setminus B_u(r) \)
• \( M \) is \( s \)-doubling: \( B \) can be covered by \( s'=(2/\delta)^{2\log s} \) balls of radius \( r'=\delta r/2 \) (invoke the doubling property multiple times)
• Consider balls with at least \( n\epsilon/3s' \) points
  – Ignore remaining points: at most \( n\epsilon/3 \) ignored
• If each ball contains a beacon: bounds are within factor \((1\pm\delta)\) of value
  • \( d_{ub}+d_{vb} \leq (d_{uv}+d_{vb}) + d_{vb} \leq 2d_{vb} \leq d_{uv}(1+\delta) \)
  • \( d_{ub}-d_{vb} \geq (d_{uv}-d_{vb})-d_{vb} \geq d_{uv}/2 \geq d_{uv}(1-\delta) \)
Proof

• If each ball contains a beacon: bounds are within factor \((1\pm\delta)\) of value
• \(\Pr[\exists b \in B_v(r') \text{ for all but a fraction of } \varepsilon/3 \text{ balls}]\) is almost 1 for a constant number of random beacons \(k=O((s'/\varepsilon)(\log(1/\varepsilon)))\)
• We ignore the remaining nodes within the \(\varepsilon/3\) fraction of balls
• This gives that all but at most an \(\varepsilon\) fraction of pairs nodes have estimations within a factor of \((1\pm\delta)\)
Lemma

- We implicitly proved:

\[ M \text{ is } s\text{-doubling, fix } \varepsilon \text{ and } \delta. \text{ Let } \\
\varepsilon' = (\varepsilon/2)(\delta/2)^{2\log s}. \text{ Then there is a set of pairs which includes all but a fraction of } \varepsilon \text{ of all pairs } (u,v) \text{ for every } u, \text{ with } \\
\min(r_u(\varepsilon), r_v(\varepsilon')) \leq \delta d_{uv} \]
Distributed Approach

• Beacons perform $O(n)$ measurements
  – A lot!
• Can nodes perform fewer measurements?
• Fully distributed algorithm
  – $\text{polylog}(n)$ for each node
Algorithm (for node u)

• Select $k$ neighbors uniformly at random and measure distances. Decide with probability $k/n$ to be a beacon.

• If u is a beacon then announce to neighbors
  – Sort measurements $(x_1, \ldots, x_k)$
  – Let $r_b = x_{2\epsilon'k}$ (estimation of $r_b(\epsilon')$)
  – Send $(b, r_b, 0)$ to all neighbors

• Upon receiving $(b, r_b, i)$ from v
  – Estimate $d_{ub}$ by $d_{uv} \pm 2ir_b$
  – If no previous message $(b, r_b, <i)$
    • and if $d_{uv} \leq 2r_b$
    • and $i < \log n$
    – then forward $(b, r_b, i+1)$ to neighbors

$\epsilon' = (\epsilon/2)(\delta'/2)^{2\log s}$
$\delta' = O(\delta/\log n)$
Algorithm (for node u)
Complexity

• Each message of beacon $b$ is forwarded at most $K = \text{clog}^2 n$ times
• $u$ broadcasts message of $b$ at most $K$ times
• $u$ receives message of $b$ at most $K$ times from each of $O(k)$ neighbors
• We take $k = O(\text{polylog} n)$

➢ Total: $\text{polylog}(n)$ work for each node
Analysis

• Estimation $d_{ub}$ by $d_{uv} \pm 2ir_b$ is valid (at least as large as real value $d_{ub}$) by induction: we got message only from node v with $d_{uv} \leq 2r_b$ whose estimation was $d_{vb} \leq 2(i-1)r_b$

• For every beacon b, by Chernoff Bounds:
  – At least $2\varepsilon'k$ neighbors with distance $r_b(4\varepsilon')$ from b
  – At most $2\varepsilon'k$ neighbors with distance $r_b(\varepsilon')$ from b

• We get: $r_b(\varepsilon') \leq r_b \leq r_b(4\varepsilon')$
Analysis

• Claim: The ball $B=B_b(r_b(\varepsilon'))$ has diameter $\leq K$ in the induced graph of random neighbors

• Proof:
  – With high probability $u$ has $\geq 3$ neighbors in $B$
    • Also for $B(r(\varepsilon'))$ of every other node
  – The graph induced by $B$ contains a constant degree expander

➢ Diameter $\leq c\log n$
Analysis

• The ball $B = B_b(r_b(\varepsilon'))$ has diameter $\leq K$ in the induced graph of random neighbors
• $w$ in $B$ will estimate $d_{wb} \leq 2r_bK$
• $u$ has a neighbor $w$ in $B$
  – With high probability
• $u$ estimates $d_{ub}$ by $d_{uw} \pm 2r_bK$
  – At most $d_{ub} \pm 3r_bK$

➢ Estimations are within $O(r_bK)$
Wrap Up

- Estimations are within $O(r_b K)$
  - We proved: M is $s$-doubling, fix $\varepsilon$ and $\delta'$. Let $\varepsilon' = (\varepsilon/2)(\delta'/2)^{2\log s}$. Then there is a set of pairs which includes all but a fraction of $\varepsilon$ of all pairs $(u, v)$ for every $u$, with $\min(r_u(\varepsilon), r_v(\varepsilon')) \leq \delta'd_{uv}$
  - W.h.p there is a beacon $b$ in $B$
  - $B \subseteq B'$: $r_b \leq r_b(4\varepsilon') \leq 2r$
  - $d_{uv}$ estimated with error $O(rK)$
    = $O(\delta d_{uv}/\log n)(c\log n)$
    = $O(\delta d_{uv})$

![Diagram showing beacon $b$, radius $2r$, and estimated error radius $r = O(\delta d_{uv}/\log n)$](image)
Next

Distance Labeling

• We are interested in **distance labeling**
  – Assign a short label to every node
  – Such that the distance between each pair of nodes can be determined by their labels
    • Or approximated by their labels
• This is useful for a **routing scheme**:  
  – According to the short labels, nodes decide how to route messages
    • Hopefully fast
  – **Stretch**: the maximum ratio (over all pairs of nodes) between the length of a routing path and the length of the shortest path
  – **Space**: size of local memory used, and routing headers attached to the packets. **Compact routing**: polylog(n) space
  – Two extremes show the tradeoff between stretch and space: full distance matrix, flooding
Distance Labeling

- We just saw one solution to distance labeling
  - The estimate gave us upper and lower bounds on the true distance by using the triangle inequality
  - But we had slack: some small fraction of pairs may not have good estimates

- Next: an algorithm for approximate distance labeling
  - With a guaranteed \((1+\delta)\)-approximation for the distances
  - Without slack
Grid Dimension

• The grid dimension of a metric is the infimum of all $\alpha$ such that for any $x \geq 2$ the cardinality of any ball is at most $x^\alpha$ times smaller than the cardinality of a ball $x$ times the radius.
  – This abstracts a useful property of d-dimensional grids (with $\alpha = d + O(1)$).

• Bounded grid dimension $\Rightarrow$ bounded doubling dimension
Assumptions

- Let \((V, d, M)\) be a metric with grid dimension \(\alpha\) and polynomially bounded aspect ratio \(\Delta\).

- Each node knows \(\alpha\) and \(\log(n \Delta)\) up to a constant factor.

- Each node has some outgoing links that collectively induce a \(d\)-degree undirected graph with expansion \(\gamma\) such that \(d/\gamma = \log^{O(1)} n\).

- Each node knows this number up to a constant factor.

- \(\delta\) in \((0, \frac{1}{2})\) is the stretch parameter.

This is basically an expander. Can be obtained whp by having each node sample a constant number of nodes as neighbors.
Main Theorem

• There exist a randomized distributed algorithm that with high probability constructs a \((1+\delta)\)-approximate triangulation of order \(O(k\log n)\). The total running time is at most \((\delta^{-\alpha} \cdot \log n)^{O(1)}\).
Rings of neighbors

• Assumption: smallest distance is 1 and largest is $\Delta$ (diameter)

• Notation: $B_{ui} = B_u(\Delta/2^i)$

• **Rings of neighbors**: a structure in which for every $1 \leq i < \log \Delta$, each node $u$ has a link to $k$ other nodes in $B_{ui}$, denoted $X_{ui}^{(i)}$ and called the $i$-th ring neighbors of $u$

• $\Delta/2^i$ is the **radius** of the ring, $k$ its **cardinality**
Rings of neighbors

- Notation: $B_{ui} = B_u(Δ/2^i)$
- **Rings of neighbors**: a structure in which for every $1 \leq i < \log Δ$, each node $u$ has a link to $k$ other nodes in $B_{ui}$, denoted $X_{ui}^{(j)}$ and called the $i$-th ring neighbors of $u$.
Randomized rings of neighbors

• A randomized algorithm for constructing rings of neighbors induces a joint probability distribution on the variables $X_{ui}^{(j)}$. 

• This distribution is called **randomized rings of neighbors (k-RRN)** if these random variables are mutually independent and distributed near-uniformly on the respective balls $B_{ui}$. 
k-RRN Algorithm

• There exists a randomized distributed algorithm that for any natural number $k$ with high probability constructs k-RRN on $(V,d_M)$ in time $O(k)(2^\alpha \cdot \log n)^{O(1)}$
  – We will not show the algorithm and proof
A directed graph $G$ is called \textit{$\delta$-zooming}, for some $\delta$ in $(0, 1)$, if for any two nodes $u, v$, node $u$ has an out-link $(u, w)$ such that $w$ lies within distance $\delta d_{uv}$ from $v$. 

\begin{itemize}
  \item A directed graph $G$ is called \textit{$\delta$-zooming}, for some $\delta$ in $(0, 1)$, if for any two nodes $u, v$, node $u$ has an out-link $(u, w)$ such that $w$ lies within distance $\delta d_{uv}$ from $v$.
\end{itemize}
δ-hierarchical node labeling

- A **δ-hierarchical node labeling** of order $k$ is a labeling of each node by one or more integers in $\{0, \ldots, \log\Delta - 1\}$ such that for each node $u$ and each label $i$ we have
  - There is at least one label-$i$ node in $B_u(r)$
  - There are at most $k$ label-$i$ nodes in $B_u(2r)$

- $r = \Delta / 2i$
k-RRN => δ-zooming and δ-hierarchical node labeling

• There exists a randomized distributed algorithm that with high probability in time $(\delta^{-\alpha} \cdot \log n)^{O(1)}$ constructs
  – A δ-zooming directed graph on $V$ of degree $\delta^{-O(\alpha)} \cdot \log^2 n$
  – A δ-hierarchical node labeling on $V$ of order $\delta^{-O(\alpha)} \cdot \log n$
Proof Idea

• Use the $k$-RRN algorithm with ring cardinality $k = \delta^{-O(\alpha)} \cdot \log n$
  – The directed graph with an edge for each node $u$ and each node $X_{ui}^{(j)}$ is a $\delta$-zooming directed graph on $V$ of degree $\delta^{-O(\alpha)} \cdot \log^2 n$, whp
  – Node $u$ assigns label $i$ to itself if and only if it is one of its own ring-$i$ neighbors. This is a $\delta$-hierarchical node labeling on $V$ of order $\delta^{-O(\alpha)} \cdot \log n$, whp
δ-hierarchical beacon network

• A δ-hierarchical beacon network of order k is a δ-hierarchical node labeling of order k such that for each node u, each label 1 ≤ i < \log \Delta, and \( r = \frac{\Delta}{2^i} \) we have
  – Node u has an i-beacon-link to some label-i nodes in \( B_u(2r) \), including all label-i nodes in \( B_u(r) \).
  – For each such beacon-link \( (u,w) \), node u knows i, the address of w, and (1+δ)-approximate estimate of \( d_M(u,w) \).

• This definition leads to a triangulation
δ-hierarchical beacon network => Triangulation

• Lemma: For any δ in (0, 1/4), any δ-hierarchical beacon network of order k is a (1 + O(δ))-approximate triangulation of order O(k log n).
• Proof:
  • The beacon set of a given node u consists of all nodes to which u has a beacon-link. For each label 1 ≤ i < logΔ there are at most k such nodes => the beacon set has size O(k log n)
  • Fix a pair (u, v) and define the distance scale as the smallest i such that r = Δ /2^i ≥ d_{uv}/(1 − δ)
  • Then there exists a label-i beacon b in B_v(δr)
  • Then d_{ub} ≤ d_{uv} + δr ≤ r, so node b is in the beacon sets of both u and v
δ-hierarchical beacon network =>
Triangulation

• Lemma: For any δ in (0, 1/4), any δ-hierarchical beacon network of order k is a (1 + O(δ))-approximate triangulation of order O(k \log n).

• Proof cont.:
• Node b is in the beacon sets of both u and v
• Let $D_u(b)$ and $D_v(b)$ be the (1+δ)-approximate upper bounds on $d_{ub}$ and $d_{vb}$, respectively

It follows that

\[ D_v(b) \leq (1+\delta)d_{vb} \leq (1+\delta)\delta r \leq (1+2\delta)d_{uv} \]
\[ D_u(b) \leq (1+\delta)d_{ub} \leq (1+\delta)(\delta r + d_{uv}) \leq (1+2\delta)d_{uv} \]
\[ D_u(b) \geq (1-2\delta)d_{uv} \]
(G,s,r,δ)-Broadcast

• A protocol in which a message is sent from s over the directed graph G and reaches all nodes in $B_s(r)$ by $(1+δ)$-stretch paths
  – And does not go too far from $B_s(r)$
Main Theorem

• There exist a randomized distributed algorithm that with high probability constructs a \( \delta \)-hierarchical beacon network of order \( k = \delta^{-o(\alpha) \cdot \log n} \), and hence a \((1+\delta)\) -approximate triangulation of order \( O(k \log n) \). The total running time is at most \((\delta^{-\alpha} \cdot \log n)^{O(1)}\)
Proof of Main Theorem

• We have a $\delta$-zooming directed graph $G^*$ with degree $k = \delta^{-O(\alpha)} \cdot \log n$ and a $\delta$-hierarchical node labeling $L$ with order $k = \delta^{-O(\alpha)} \cdot \log n$.

• Each node $u$ starts a $(G,u,r,\delta)$-broadcast for $r = \Delta / 2^i$

• Each node $w$ in $B_u(r)$ receives this broadcast and gets a $(1+\delta)$-approximation estimate of $d_{uw}$

=> We constructed a $\delta$-hierarchical beacon network of order $k$
Proof of Main Theorem cont.

- $L_{uv}^i$ is the load on $u$ by the $(G,u,r,\delta)$-broadcast
- $L_{uv}^i = 0$ for every node $v$ not labeled $i$
- There are at most $k$ label-$i$ nodes in $B_u(2r)$
- $S_{ui} = \text{the set of all such nodes}$
- $L_{uv}^i = O(k)$ if $v$ is in $S_{ui}$, and $L_{uv}^i = 0$ otherwise
- The load on $u$ is therefore

\[ \Sigma_{vi} L_{uv}^i \leq \Sigma_i O(k) |S_{ui}| \leq O(k^2 \log n) \]