# Distance Estimation and Object Location via Rings of Neighbors

Alex Slivkins

Cornell University

# Four problems, one technique

- routing schemes:  look at routing table, choose next hop
- small-world networks:  Alice efficiently finds Bob using social links, by zooming in with respect to some notion of distance
- distance labeling schemes: decode distances from short node labels
- triangulation: each node stores distances to small #other nodes; reconstruct distances via triangle inequality

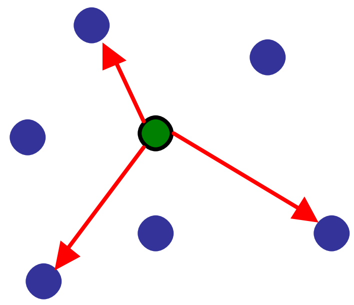
What's common in all four problems?

- underlying notion of distance (a given metric)
- we need to encode distances and paths via short node labels

We use a unified technique called "rings of neighbors":
overlay network with multiple scales of resolution

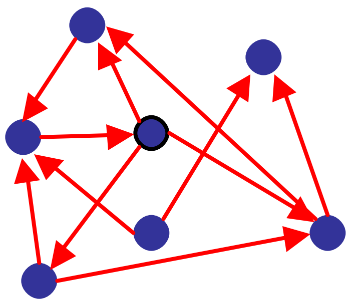
# “Network of neighbors”

- we are given a distance function d(u,v) on nodes
- each node stores distances to some other nodes: its neighbors
  - some neighbors can be far, some can be close



# "Network of neighbors"

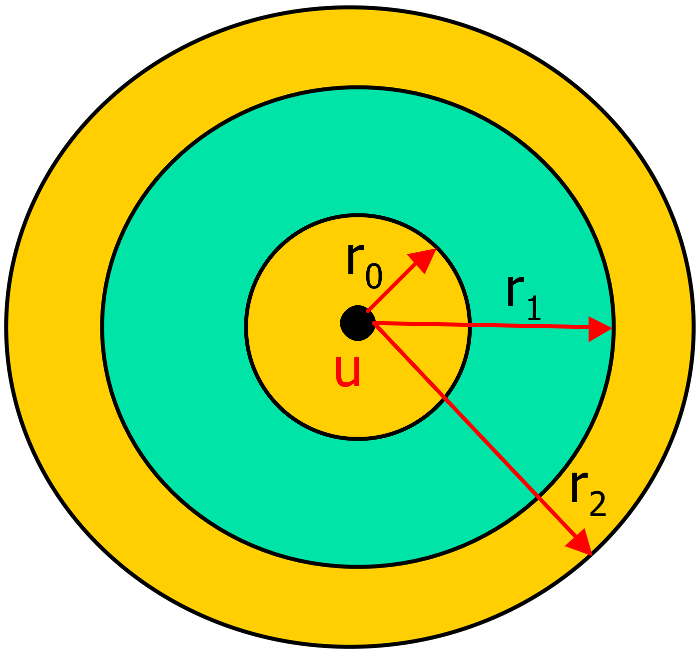
- we are given a distance function d(u,v) on nodes
- each node stores distances to some other nodes: its neighbors
  - some neighbors can be far, some can be close
- neighbors induce a a directed graph: network of neighbors



# "Rings of neighbors"

Neighbors of each node u are organized in rings

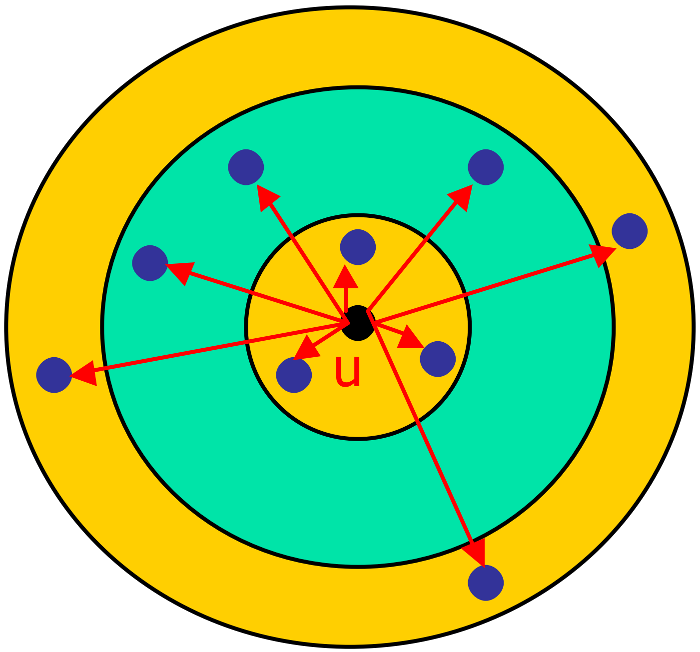
- sequence of radii $r_0(u) < r_1(u) < r_2(u) < ...$



# “Rings of neighbors”

Neighbors of each node u are organized in rings

- sequence of radii $r_0(u) < r_1(u) < r_2(u) < ...$
- ring #i: k neighbors distributed [somehow] in $B(u, r_i(u))$




- usually $k = O(\log n)$
- radii $r_i(u)$ can depend on u
- the choice of radii and the choice of neighbors depends on the problem

# Doubling metrics

- Def any ball can be covered by $2^{\alpha}$ balls of half the radius, $\alpha$ = const
  - abstracts the doubling property of low-dim Euclidean metrics: for any point set in a $k$-dim Euclidean metric, $\alpha = O(k)$
  - doubling dimension = smallest such $\alpha$
- Why doubling metrics are interesting?
  - subsume and extend const-dim Euclidean metrics and growth-constrained metrics
  - in many recent TCS papers, lead to efficient constructions
  - clean model for distances that are based in geography, but perturbed by various smaller factors

# Plan

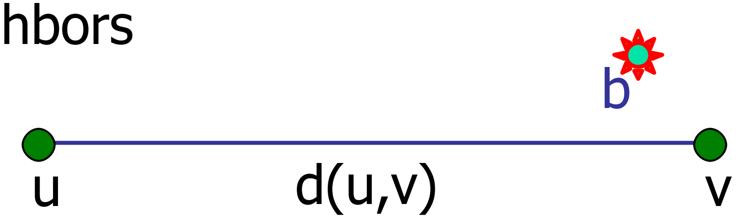
- ☑ Introduction
- ☐ Triangulation and distance labeling schemes
- ☐ One clean proof (for triangulation)
- ☐ Extensions

# Distance labeling schemes

- In a Distance Labeling Scheme for metric (V,d)
  - each node u is assigned a label $L_u$
  - d(u,v) is estimated by a function $f(L_u, L_v)$
  - $(1+\delta)$-approximate if $d(u,v) \le f(L_u, L_v) \le (1+\delta)\, d(u,v)$
- DLS is a compact and distributed data structure
  - encodes $n \times n$ matrix of distances as $n \times 1$ array of labels
- Goal: for a given metric and $\delta$, minimize label length
  - trivial DLS : $L_u$ consists of distances from u to all other nodes

# Triangulation: a specific kind of DLS

- Triangulation = DLS where label of each node u consists of distances to k other nodes ("neighbors" of u)
- To estimate d(u,v), use triangle inequality on triples (u,v,b), $b \in S$ = {common neighbors of u and v} :
  - one triple: $|d(u,b)-d(v,b)| \leq d(u,v) \leq d(u,b)+d(v,b)$
  - upper bound: $D^{+}(u,v) = \min_{b \in S} d(u,b)+d(v,b)$
  - lower bound: $D^{-}(u,v) = \max_{b \in S} |d(u,b)-d(v,b)|$
- Given $\delta>0$, ensure $D^{+}/D^{-} \leq 1+\delta$, minimize k = #neighbors



# Triangulation: networking motivation

- triangulation as distributed algorithm in peer-to-peer network
  - all nodes have the same set of neighbors (beacons) that are selected uniformly at random in the network
  - defined in [Kleinberg, Slivkins, Wexler FOCS'04], motivated by a systems project: IDMaps [Francis+ INFOCOM '01]
  - theoretical guarantees: $D^+ / D^- \leq 1+\delta$ for most node pairs; motivated by empirical performance of IDMaps.
- other systems that estimate distances via triangle inequality [Hotz '94, Guyton+'95, KSB'01], [Wong, Slivkins, Sirer SIGCOMM '05]

# Some results

Thm For each $\delta>0$ there exists (and can be efficiently constructed) a triangulation such that $D^{+}/D^{-} \leq 1+\delta$ for all edges and each node has $C \log(n)$ neighbors, $C = (1/\delta)^{O(\text{doubling dim})}$

- From triangulation to a distance labeling scheme
  - in label(u), for each neighbor of u store id and distance
  - use (log n)-bit ids, store distances as log($1/\delta$)-bit mantissa and (log log T)-bit exponent, where all distances lie in [1, T]
  - Total: C (log n) ((~~log n~~)+(log log T))-bit labels

bells & whistles

# Our result in context

- $(1+\delta)$-approximate distance labeling scheme with $k$-bit labels
  - $k = \Omega(n^{2/3})$ for general metrics [Gavoille+ SODA '01]
- Much better for doubling metrics:
  - assume all distances lie in $[1, T]$; let $C = (1/\delta)^{O(\text{doubling dim})}$
  - $k = C \log(n) \log(T)$ [Gupta+ FOCS '03]
  - $k = C \log(T)$ [Talwar STOC '04]
- what if $T = 2^n$ ? Can we reduce the dependency n $T$?
  - $k = C \log(n) ((\log n) + (\log \log T))$ } ngulation
  - $k = C \log (n) (\log \log T)$ bells & whistles
  - matching lower bound [Mendel, Har-Peled SoCG '05]

alternative proof in [Mendel, Har-Peled'05]

# Plan

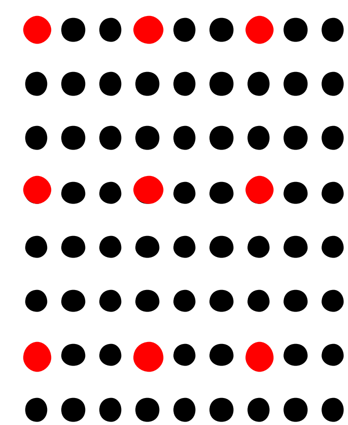
- ☑ Introduction
- ☑ Triangulation and distance labeling schemes
- ☐ One clean proof (for triangulation)
- ☐ Extensions

# Triangulation: proof

Thm  For each $\delta>0$ there exists (and can be efficiently constructed) a triangulation such that $D^+/D^- \leq 1+\delta$ for all edges and each node has $C \log(n)$ neighbors, $C = (1/\delta)^{O(\text{doubling dim})}$

Def: Set N of nodes is a t-net for metric (V,d) if

- any node in V is within distance at most t from N
- distance between any two nodes in N is at least t



Fact: t-net exists and can be constructed greedily;

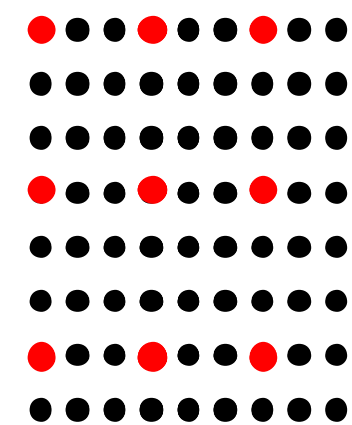
In a doubling metric, any ball of radius O(t) contains at most a constant #points from N

# Triangulation: proof

Thm For each $\delta>0$ there exists (and can be efficiently constructed) a triangulation such that $D^+/D^- \leq 1+\delta$ for all edges and each node has $C \log^2(n)$ neighbors, $C = (1/\delta)^{O(\text{doubling dim})}$

Def: Set $N$ of nodes is a $t$-net for metric $(V,d)$ if

- any node in $V$ is within distance at most $t$ from $N$
- distance between any two nodes in $N$ is at least $t$

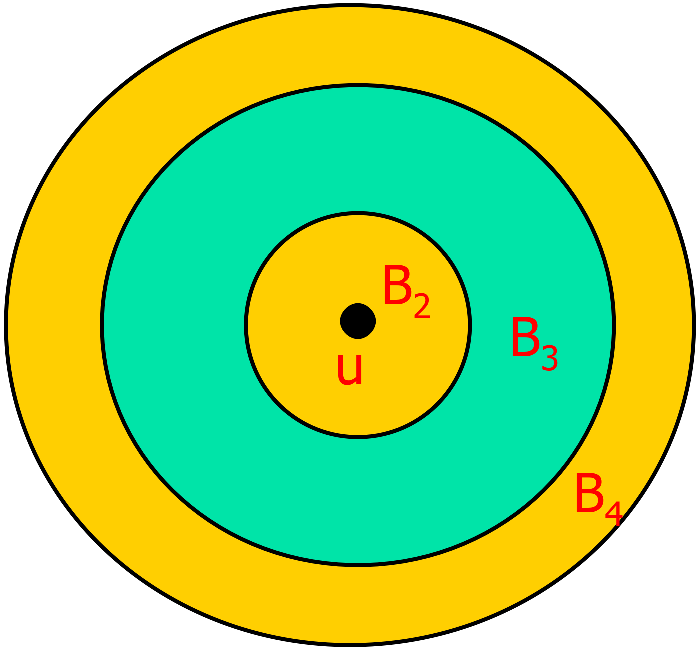


Fact: $t$-net exists and can be constructed greedily;

In a doubling metric, any ball of radius $O(t)$ contains at most a constant #points from $N$

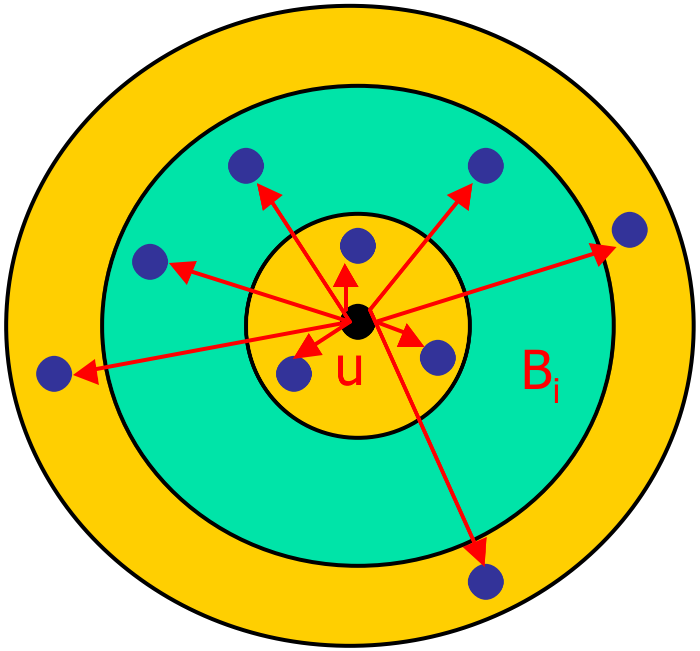
# Construction: neighbors of node u

- $B_i$ = smallest ball around u that contains $2^i$ nodes
- Use "rings of neighbors": all ring i neighbors lie in $B_i$



# Construction: neighbors of node u

- $B_i$ = smallest ball around u that contains $2^i$ nodes
- Use "rings of neighbors": all ring i neighbors lie in $B_i$

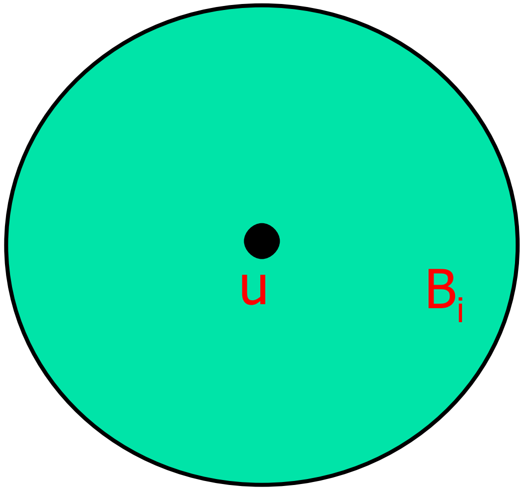


# Construction: neighbors of node u

- $B_i$ = smallest ball around u that contains $2^i$ nodes
- Use "rings of neighbors": all ring i neighbors lie in $B_i$

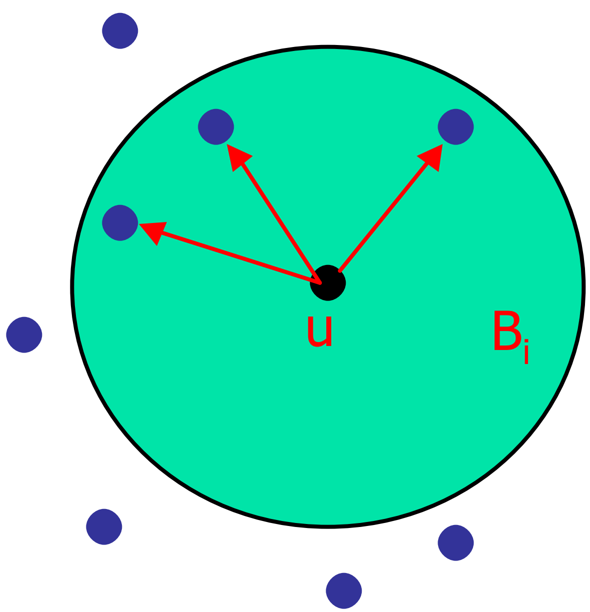
Two kinds of neighbors. First kind:

- select with respect to cardinality



# Construction: neighbors of node u

- $B_i$ = smallest ball around u that contains $2^i$ nodes
- Use "rings of neighbors": all ring i neighbors lie in $B_i$



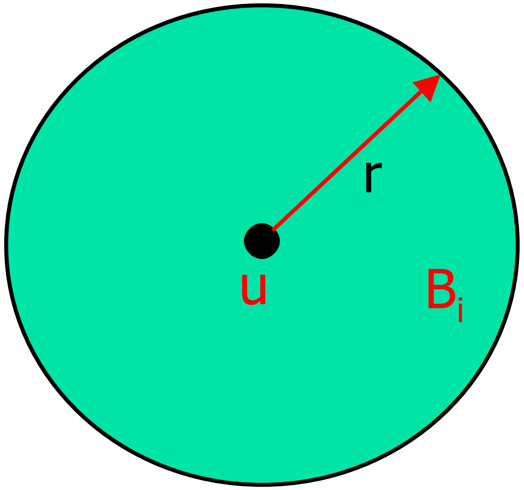

Two kinds of neighbors. First kind:

- select with respect to cardinality
- pick random $\Theta(\log n)(n/2^i)$-node set S
  - blue dots = elements of S
  - one such set for all nodes u
- ring i neighbors of u are nodes in $B_i \cap S$
- $\Theta(\log n)$ ring i neighbors
  in expectation and with high probability

# Construction: neighbors of node u

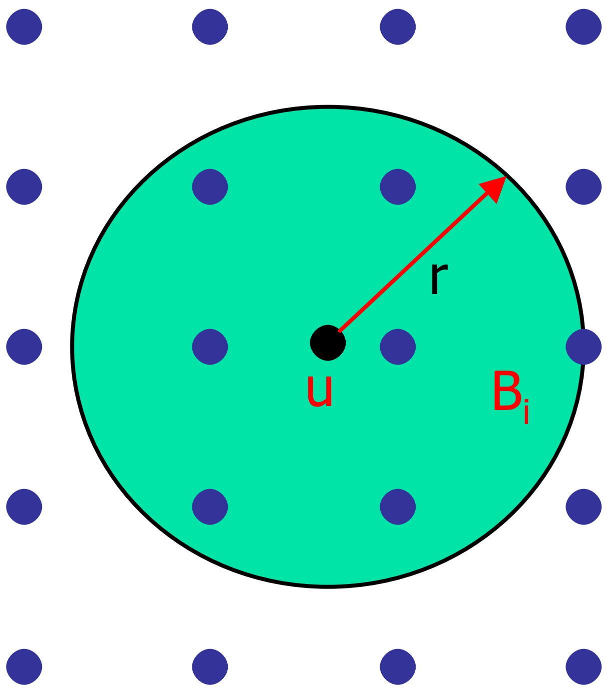
- $B_i$ = smallest ball around u that contains $2^i$ nodes
- Use “rings of neighbors”: all ring i neighbors lie in $B_i$

Second kind of neighbors: use t–nets



# Construction: neighbors of node u

- $B_i$ = smallest ball around u that contains $2^i$ nodes
- Use "rings of neighbors": all ring i neighbors lie in $B_i$

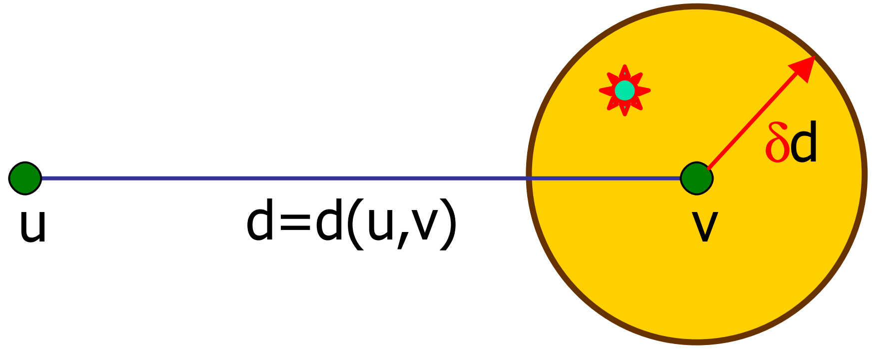



Second kind of neighbors: use t–nets

- let t = smallest $2^j$ that is at least $\delta r$
- fix some t–net N (one t–net for all u)
  - recall: $d(\text{any node}, N) \le t$
    $d(v, v') \ge t$ for any $v, v' \in N$
- ring i neighbors of u are nodes in N that lie within distance $r/\delta$ from u
- doubling metric $\Rightarrow$ at most $(1/\delta)^{O(\text{doubling dim})}$ such neighbors

# Proof of correctness (sketch)

- want: a common neighbor within distance δd(u,v) from u or v

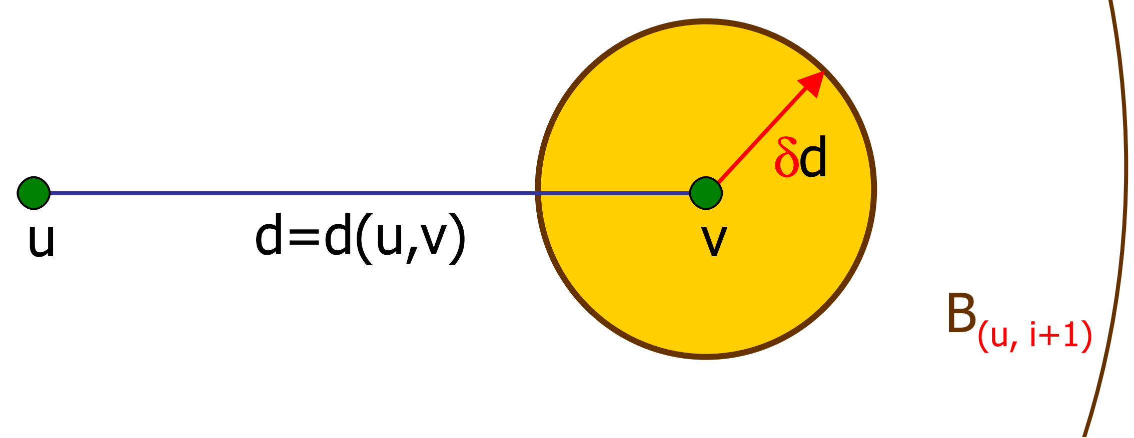


# Proof of correctness (sketch)

- want: a common neighbor within distance $\delta d(u,v)$ from u or v
- $B_{(u,i)}$ = smallest ball around u that contains $2^i$ nodes

---

- pick smallest i such that $B(v,\delta d) \subset B_{(u,\,i+1)}$

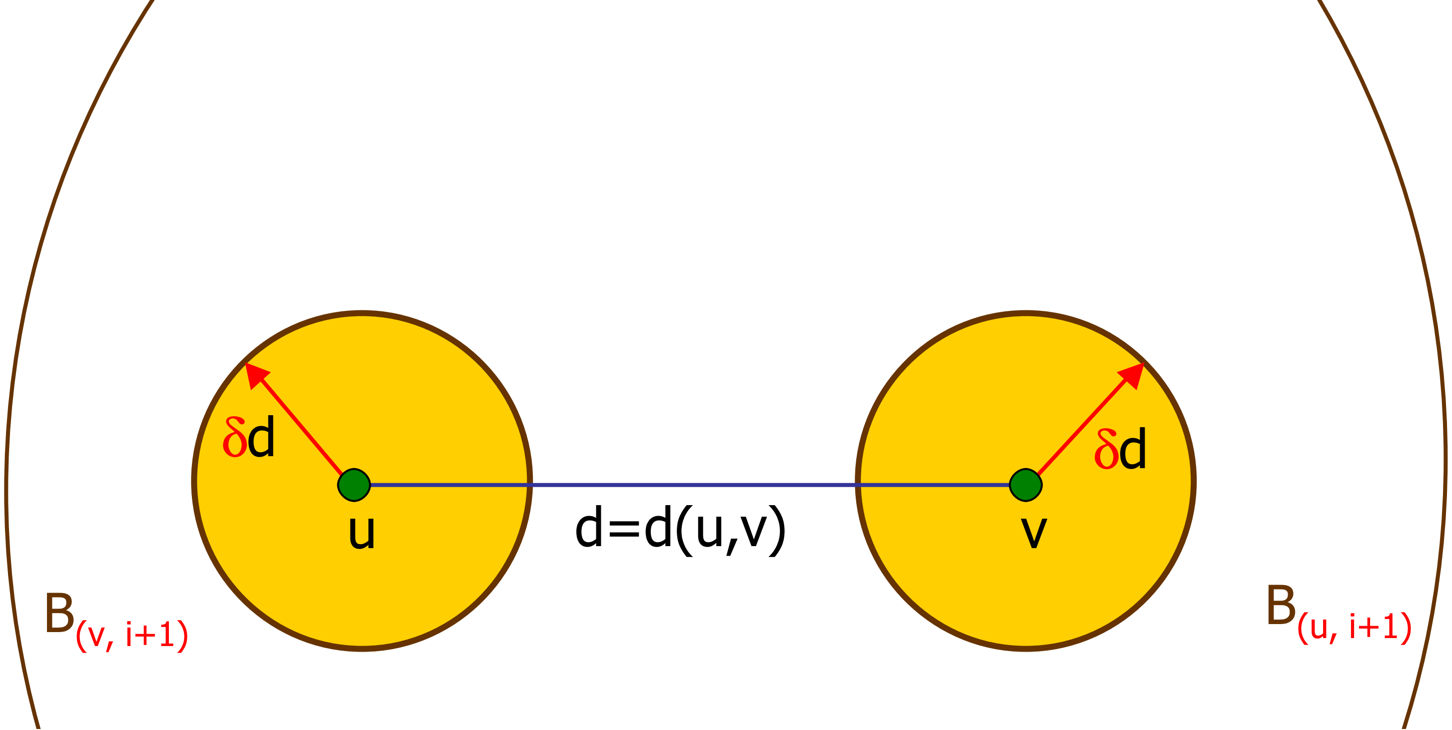


# Proof of correctness (sketch)

- want: a common neighbor within distance $\delta d(u,v)$ from u or v
- $B_{(u,i)}$ = smallest ball around u that contains $2^i$ nodes

---

- pick smallest i such that $B(v,\delta d) \subset B_{(u,\,i+1)}$ and $B(u,\delta d) \subset B_{(v,\,i+1)}$

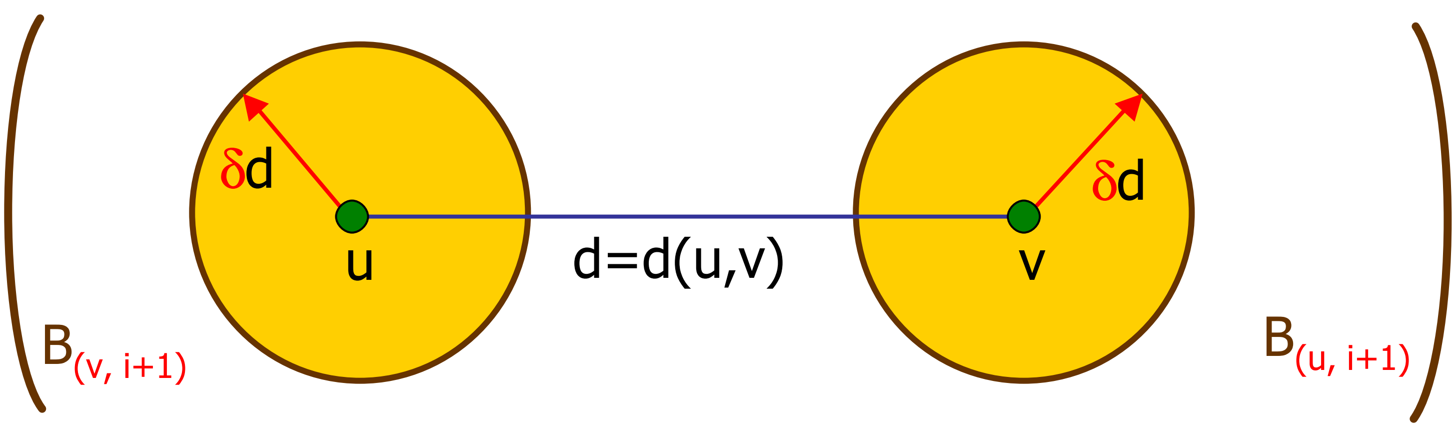


# Proof of correctness (sketch)

- want: a common neighbor within distance $\delta d(u,v)$ from u or v
- $B_{(u,i)}$ = smallest ball around u that contains $2^i$ nodes

---

- pick smallest i such that $B(v,\delta d) \subset B_{(u, i+1)}$ and $B(u,\delta d) \subset B_{(v, i+1)}$

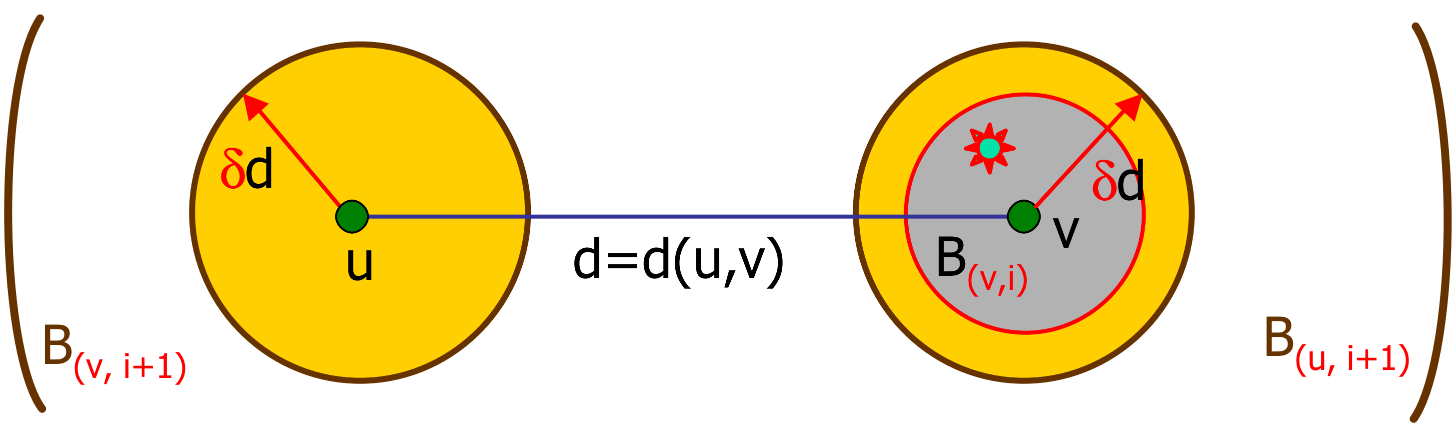


# Proof of correctness (sketch)

- want: a common neighbor within distance $\delta d(u,v)$ from u or v
- $B_{(u,i)}$ = smallest ball around u that contains $2^i$ nodes

---

- if $B_{(v,i)}$ has radius at most $\delta d$ then we are done!
  - $B_{(v,i)}$ contains a common neighbor of the first kind
    (recall: such neighbors are selected with respect to cardinality)

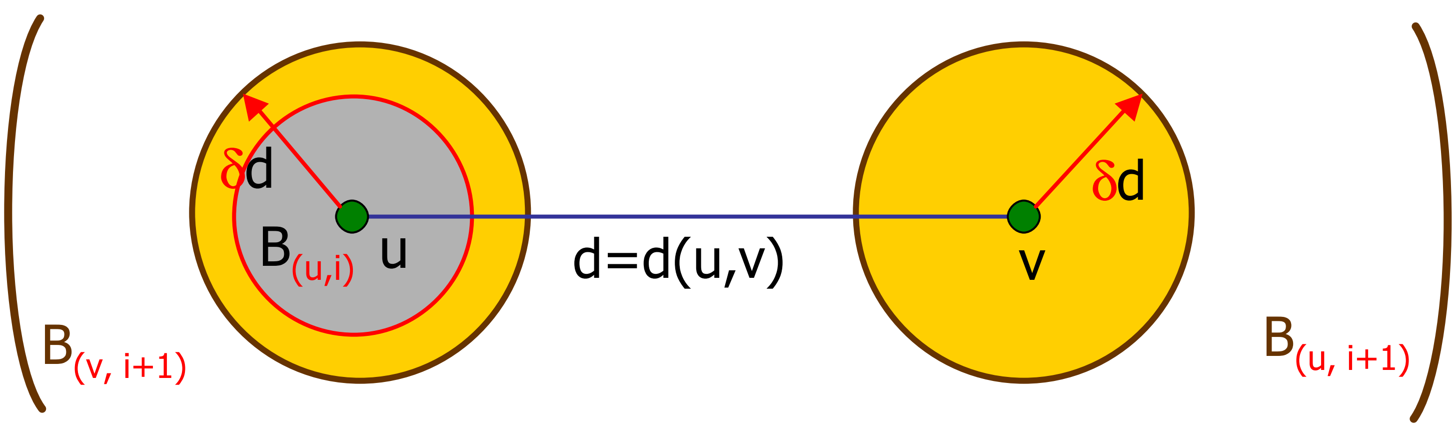


# Proof of correctness (sketch)

- want: a common neighbor within distance $\delta d(u,v)$ from u or v
- $B_{(u,i)}$ = smallest ball around u that contains $2^i$ nodes

---

- if $B_{(v,i)}$ has radius at most $\delta d$ then we are done!
  - $B_{(v,i)}$ contains a common neighbor of the first kind (recall: such neighbors are selected with respect to cardinality)
- Similarly, we are done if $B_{(u,i)}$ has radius at most $\delta d$

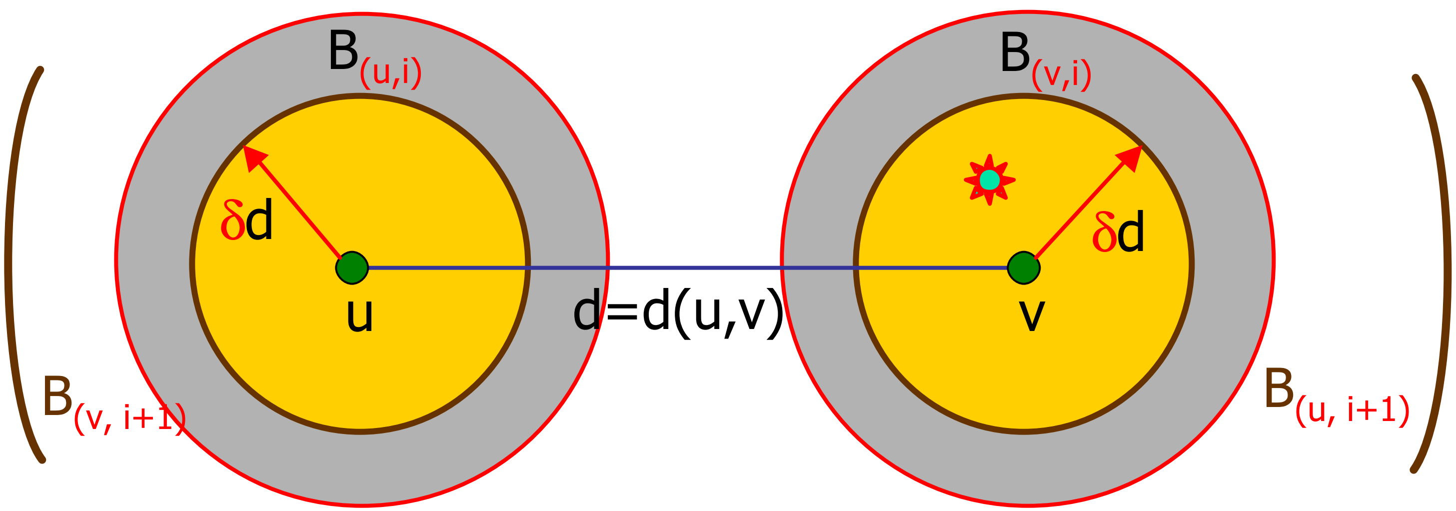


# Proof of correctness (sketch)

- want: a common neighbor within distance $\delta d(u,v)$ from u or v
- $B_{(u,i)}$ = smallest ball around u that contains $2^i$ nodes

---

- now assume both $B_{(u,i)}$ and $B_{(v,i)}$ have radii $> \delta d$
- then $B(v,\delta d)$ contains a common neighbor of the second kind
  (recall: such neighbors are selected from t-nets)



# Plan

- ☑ Introduction
- ☑ Triangulation and distance labeling schemes
- ☑ One clean proof (for triangulation)
- ☐ Extensions

# Extensions

In this paper: apply "rings of neighbors" to design improved routing schemes and small-world networks

Further work: "rings of neighbors" as an underlying layer for distributed networking applications

- system for nearest neighbor queries in peer-to-peer networks
  - [Wong, Slivkins, Sirer SIGCOMM'05]
- other applications, e.g. diverse node set, multi-constraint selection
- open questions, both theoretical and practical
  - what properties should "rings of neighbors" have? how to construct and maintain them?