Separating Hierarchical and General Hub Labelings

Andrew V. Goldberg[1] Ilya Razenshteyn[2] Ruslan Savchenko[3]

[1]Microsoft Research Silicon Valley
[2]CSAIL, MIT
[3]Moscow State University

June 23, 2013

Overview

Preliminaries

Tight Bound for HHL

An $O(2.83^d)$ HL

HL Bound

Shortest Paths with Preprocessing

Definition (Shortest Path Problem)

Given a graph and two vertices find a shortest path between them. Sometimes only distance is required.

Definition (Shortest Paths with Preprocessing)

At preprocessing time algorithm prepares some data. In query time algorithm uses the prepared data.

Shortest paths with preprocessing may give sublinear query time.

Labeling Approach

Definition (Labels, Journal of Graph Theory'00 D. Peleg)

- For each vertex v, preprocessing computes a label $L(v)$.
- The u, v distance is computed only from $L(u)$ and $L(v)$ (without using the graph).

Definition (Hub Labels (HL), SODA'02 Cohen et al.)

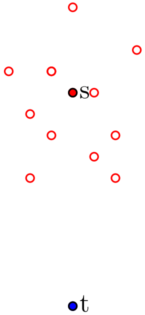
- Label $L(v)$ is a set of pairs $(u, dist(v, u))$.
- Cover property: $\forall u, v,\ L(u) \cap L(v)$ contains a vertex on the shortest $u - v$ path.

Hub Labels Query

- Query s, t distance

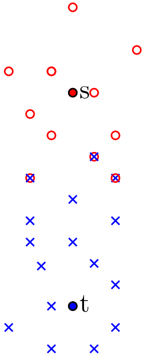
Hub Labels Query

- Query s, t distance
- Look at $L(s)$



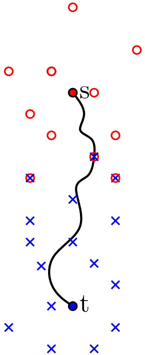
Hub Labels Query

- Query s, t distance
- Look at $L(s)$
- Look at $L(t)$



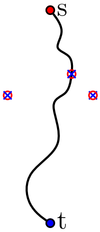
Hub Labels Query

- Query s, t distance
- Look at $L(s)$
- Look at $L(t)$
- By definition shortest s, t path contains common hub



Hub Labels Query

- Query s, t distance
- Look at $L(s)$
- Look at $L(t)$
- By definition shortest s, t path contains common hub
- Look at $L(s) \cap L(t)$

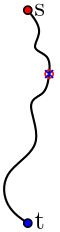


Hub Labels Query

- Query s, t distance
- Look at $L(s)$
- Look at $L(t)$
- By definition shortest s, t path contains common hub
- Look at $L(s) \cap L(t)$
- Common hub v with the smallest $dist(s, v) + dist(v, t)$ is the right one

Hub Labels Query

- Query s, t distance
- Look at $L(s)$
- Look at $L(t)$
- By definition shortest s, t path contains common hub
- Look at $L(s) \cap L(t)$
- Common hub v with the smallest $dist(s, v) + dist(v, t)$ is the right one
- Query time is $O(|L(s)| + |L(t)|)$.



General and Hierarchical Hub Labels

- SODA'02 Cohen et al.
  $O(\log n)$ approximation for the total size of labels by reduction to the set cover.
  $O(n^4)$ time.
- ESA'12 Abraham et al.
  Hierarchical labels: there is an order on vertices and each vertex has only more important nodes in its label.
  $O^*(nm)$ time.

General and Hierarchical Hub Labels

- SODA'02 Cohen et al.
  $O(\log n)$ approximation for the total size of labels by reduction to the set cover.
  $O(n^4)$ time.
- ESA'12 Abraham et al.
  Hierarchical labels: there is an order on vertices and each vertex has only more important nodes in its label.
  $O^*(nm)$ time.
- Hierarchical labels are easier to compute, but is the label size good compared to general ones?

General and Hierarchical Hub Labels

- SODA'02 Cohen et al.
  $O(\log n)$ approximation for the total size of labels by reduction to the set cover.
  $O(n^4)$ time.
- ESA'12 Abraham et al.
  Hierarchical labels: there is an order on vertices and each vertex has only more important nodes in its label.
  $O^*(nm)$ time.
- Hierarchical labels are easier to compute, but is the label size good compared to general ones?

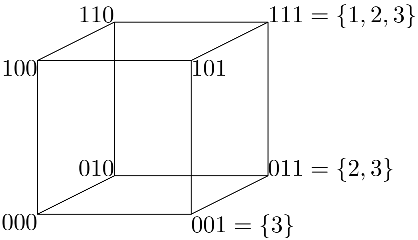
Goal: compare general (HL) vs hierarchical (HHL) hub labels.

Hypercubes

Definition (Hypercube)

A graph H with vertex set $\{0,1\}^d$ where vertices are adjacent iff their IDs differ in exactly one bit is called *(binary) hypercube*.

Sometimes we interpret vertex IDs as subsets of $\{1,\ldots,d\}$.



Our Contribution

We present tight bounds for both HHL and HL on hypercubes.
It follows that that HHL can be polynomially bigger than HL.

Overview

Preliminaries

Tight Bound for HHL

An $O(2.83^d)$ HL

HL Bound

An HHL of size 3^d

- Treat vertex IDs as subsets of $\{1, \ldots, d\}$.
- A vertex w is in $L(v)$ iff $w \subseteq v$.
- For u, v we have $u \cap v$ both in $L(u)$ and $L(v)$ and it is on the shortest $u - v$ path.
- Labeling is HHL (order vertices by decreasing their IDs).
- Labeling size is

$$\sum_{i=0}^{d} 2^i \binom{d}{i} = 3^d.$$

Canonical HHL

- For v, w the *induced hypercube* H_{vw} is the subgraph induced by vertices on all $v - w$ shortest paths.

$$\begin{array}{l} v = 01001\,1001\,10 \\ w = 01001\underbrace{0111}_{\text{arbitrary in } H_{vw}}10 \end{array}$$

- For a fixed order of vertices $v_1, v_2, \ldots, v_n$, define a *canonical labeling*: w is in $L(v)$ iff w is the maximum vertex of H_{vw}.
- This is the minimum HHL with respect to the vertex order.
- The previous labeling of size 3^d is canonical.

Canonical HHL Size Independence

Theorem

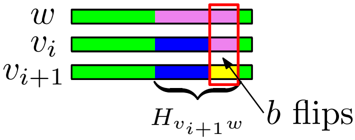
The size of a canonical labeling is independent of the vertex order.

Proof:

- Suppose we transpose v_i and v_{i+1}.
- $L(w)$ changes only if $v_i \in H_{v_{i+1}w}$ or $v_{i+1} \in H_{v_i w}$, and v_{i+1} is the most important in $H_{v_{i+1}w}$ or $H_{v_i w}$ respectively.
- Consider $b : H \mapsto H$, b flips bits in which v_i and v_{i+1} differ.
- Claim: v_{i+1} is removed from $L(w)$ iff v_i is added to $L(b(w))$.

Canonical HHL Size Independence

- If v_{i+1} is removed from $L(w)$, then $v_i \in H_{v_{i+1}w}$ and v_{i+1} is the maximum in $H_{v_{i+1}w}$.
- From $v_i \in H_{v_{i+1}w}$: v_i coincides with v_{i+1} in the positions in which v_{i+1} and w coincide. So in the latter positions v_{i+1}, w, $b(v_{i+1})$ and $b(w)$ coincide (since b doesn't touch them).



- We have $H_{v_{i+1}w} = H_{b(v_{i+1})b(w)} = H_{v_i b(w)}$. First, v_{i+1} is the maximum vertex, then it is v_i. So v_i is added to $L(b(w))$.

□

Overview

Preliminaries

Tight Bound for HHL

An $O(2.83^d)$ HL

HL Bound

Simple HL for Hypercubes

- Labeling: $w \in L(v)$ iff the first $\lfloor d/2 \rfloor$ or last $\lceil d/2 \rceil$ bits of w are identical to those of v.
- For s,t common hub has first bits from s and last from t.
- It is non-hierarchical since $w \in L(v)$ implies $v \in L(w)$.
- Its size is $O(2^{\frac{3}{2}d}) = O(2.83^d)$.

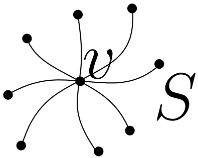
Simple HL for Hypercubes

- Labeling: $w \in L(v)$ iff the first $\lfloor d/2 \rfloor$ or last $\lceil d/2 \rceil$ bits of w are identical to those of v.
- For s,t common hub has first bits from s and last from t.
- It is non-hierarchical since $w \in L(v)$ implies $v \in L(w)$.
- Its size is $O(2^{\frac{3}{2}d}) = O(2.83^d)$.
- This implies that optimal HL for hypercubes are polynomially smaller than HHL.

Overview

Labeling as a Set Cover

- The labeling problem is a special case of **SET-COVER** (SODA'02 Cohen et al.).
- Unordered pairs of vertices are covered by packs of shortes paths with a common hub.
- Cost is the number of endpoints.



ILP Formulation

- For a vertex v and a subset of vertices S variable $x_{v,S}$ equals to one iff S is the set of vertices whose labels contain v.
- ILP for labeling:

$$\min \sum_{v,S} |S| \cdot x_{v,S} \qquad \text{subject to}$$

$$\begin{cases} x_{v,S} \in \{0,1\} & \forall\, v \in \{0,1\}^d, S \subseteq \{0,1\}^d \\ \sum_{\substack{S \supseteq \{i,j\} \\ v \in H_{ij}}} x_{v,S} \geq 1 & \forall \{i,j\} \subseteq \{0,1\}^d \end{cases} \tag{ILP}$$

- We denote the optimal value of (ILP) by OPT.

LP Relaxation

- LP-relaxation of (ILP):

$$\min \sum_{v,S} |S| \cdot x_{v,S} \qquad \text{subject to}$$

$$\begin{cases} x_{v,S} \geq 0 & \forall\, v \in \{0,1\}^d, S \subseteq \{0,1\}^d \\ \sum_{\substack{S \supseteq \{i,j\} \\ v \in H_{ij}}} x_{v,S} \geq 1 & \forall \{i,j\} \subseteq \{0,1\}^d \end{cases} \tag{P}$$

- We denote the optimal value of (P) by LOPT,
- It is known that $\mathrm{LOPT} \leq \mathrm{OPT} \leq O(d) \cdot \mathrm{LOPT}$.

Dual LP

- Dual program to (P).

$$\max \sum_{\{i,j\}} y_{\{i,j\}} \qquad \text{subject to}$$

$$\begin{cases} y_{\{i,j\}} \geq 0 & \forall \{i,j\} \subseteq \{0,1\}^d \\ \sum_{\substack{\{i,j\} \subseteq S \\ H_{ij} \ni v}} y_{\{i,j\}} \leq |S| & \forall v \in \{0,1\}^d, S \subseteq \{0,1\}^d \end{cases} \tag{DP}$$

- The dual problem is a packing problem.
- LOPT is also the optimal solution value for (DP).

Regular Dual LP

- We require the values $y_{\{i,j\}}$ depend only on the $\mathrm{dist}(i,j)$.
- We get variables $\tilde{y}_0, \tilde{y}_1, \ldots, \tilde{y}_d$.
- Let N_k denote the number of vertex pairs at distance k.
- *Regular* program:

$$\max \sum_k N_k \cdot \tilde{y}_k \quad \text{subject to}$$
$$\begin{cases} \tilde{y}_k \geq 0 & \forall\, 0 \leq k \leq d \\ \sum_{\substack{\{i,j\} \subseteq S \\ H_{ij} \ni 0^d}} \tilde{y}_{\mathrm{dist}(i,j)} \leq |S| & \forall S \subseteq \{0,1\}^d \end{cases} \tag{RP}$$

- Denote optimal value by ROPT. Clearly $\mathrm{ROPT} \leq \mathrm{LOPT}$.

ROPT $\geq$ LOPT

Lemma

ROPT $\geq$ LOPT.

To prove it we show that if $y_{\{i,j\}}$ is feasible for (DP) then

$$\tilde{y}_k = \frac{\sum_{\{i,j\}:\mathrm{dist}(i,j)=k} y_{\{i,j\}}}{N_k}$$

is feasible for (RP).

Reduction to single $\tilde{y}_k$

- So far we've seen that:
  - $\mathrm{LOPT} \leq \mathrm{OPT} \leq O(d) \cdot \mathrm{LOPT}$
  - $\mathrm{LOPT} = \mathrm{ROPT}$

Reduction to single $\tilde{y}_k$

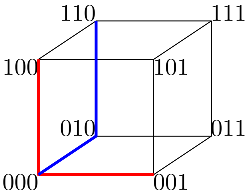
- So far we've seen that:
  - $\mathrm{LOPT} \le \mathrm{OPT} \le O(d) \cdot \mathrm{LOPT}$
  - $\mathrm{LOPT} = \mathrm{ROPT}$
- Let $\tilde{y}_k^*$ denote the maximum feasible value of $\tilde{y}_k$. Since $\mathrm{ROPT} = \max \sum_k N_k \cdot \tilde{y}_k$, it's clear that $\max_k N_k \tilde{y}_k^* \le \mathrm{ROPT} \le (d+1) \cdot \max_k N_k \tilde{y}_k^*$.

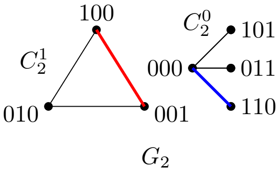
Reduction to single $\tilde{y}_k$

- So far we've seen that:
  - $\mathrm{LOPT} \leq \mathrm{OPT} \leq O(d) \cdot \mathrm{LOPT}$
  - $\mathrm{LOPT} = \mathrm{ROPT}$
- Let $\tilde{y}_k^*$ denote the maximum feasible value of $\tilde{y}_k$. Since $\mathrm{ROPT} = \max \sum_k N_k \cdot \tilde{y}_k$, it's clear that $\max_k N_k \tilde{y}_k^* \leq \mathrm{ROPT} \leq (d+1) \cdot \max_k N_k \tilde{y}_k^*$.
- Next we find $\tilde{y}_k^*$ and then $\max_k N_k \tilde{y}_k^*$.

Central Graphs G_k

- Define graphs G_k: vertices are the same, two vertices are adjacent iff there is a shortest path of length k between them that passes through 0^d.
- For any subgraph (S, E) of G_k, $\tilde{y}_k^*$ is smaller than $|S|/|E|$, since $\tilde{y}_k \cdot |E| = \sum_{\{i,j\}\in E} \tilde{y}_k \leq \sum_{\{i,j\}\subseteq S, H_{ij}\ni 0^d} \tilde{y}_{\mathrm{dist}(i,j)} \leq |S|$.
- Let C_k^i denote the component with sets of cardinality i (here vertices are interpreted as subsets of $\{1, \ldots, d\}$).
- C_k^i density is $\binom{d}{i} \cdot \binom{d-i}{k-i} / (\binom{d}{i} + \binom{d}{k-i})$.





Regular Graph Densest Subgraph

Lemma

In a regular graph, density of any subgraph does not exceed the density of the graph.
In a regular bipartite graph (i.e., degrees of each part are uniform), the density of any subgraph does not exceed the density of the graph.

Max Density of G_k

Lemma

For fixed d and k with $k \leq d$, the minimum of the expression $\left(\binom{d}{x} + \binom{d}{k-x}\right) / \binom{d}{x} \cdot \binom{d-x}{k-x}$ is achieved for $x = \lfloor k/2 \rfloor$ and $x = \lceil k/2 \rceil$ (with the two values being equal).

Final Steps

We have

$$\tilde{y}_k^* = \begin{cases} 1 & k = 0 \\ 2/\binom{d-i}{i} & k = 2i, i > 0 \\ \left(\binom{d}{i} + \binom{d}{i+1}\right) / \left(\binom{d}{i} \cdot \binom{d-i}{i+1}\right) & k = 2i + 1. \end{cases}$$

Now we need to find the maximum value of $\psi(k) := N_k \cdot \tilde{y}_k^* =$

$$= 2^d \cdot \begin{cases} \binom{d}{2i}/\binom{d-i}{i} & k = 2i \\ \binom{d}{2i+1} \cdot \left(\binom{d}{i} + \binom{d}{i+1}\right) / \left(2 \cdot \binom{d}{i} \cdot \binom{d-i}{i+1}\right) & k = 2i + 1. \end{cases}$$

Final Steps

We have

$$\tilde{y}_k^* = \begin{cases} 1 & k = 0 \\ 2/\binom{d-i}{i} & k = 2i, i > 0 \\ \left(\binom{d}{i} + \binom{d}{i+1}\right) / \left(\binom{d}{i} \cdot \binom{d-i}{i+1}\right) & k = 2i + 1. \end{cases}$$

Now we need to find the maximum value of $\psi(k) := N_k \cdot \tilde{y}_k^* =$

$$= 2^d \cdot \begin{cases} \binom{d}{2i}/\binom{d-i}{i} & k = 2i \\ \binom{d}{2i+1} \cdot \left(\binom{d}{i} + \binom{d}{i+1}\right) / \left(2 \cdot \binom{d}{i} \cdot \binom{d-i}{i+1}\right) & k = 2i + 1. \end{cases}$$

We show $\max_k N_k \cdot \tilde{y}_k^* = (2.5 + o(1))^d$.

HL Bound

Theorem

Optimal value for HL *on* d*-dimensional hypercube is* $(2.5 + o(1))^d$.

Concluding Remarks

- We show a polynomial gap between the sizes of HL and HHL on hypercubes.
- Our proof for $(2.5 + o(1))^d$-size HL is non-constructive, but Cohen et al. algorithm can build such labels.
- Explicit $(2.5 + o(1))^d$-size HL construction is an open question.

Thank You!

Questions?