Distance Labeling in Graphs

Cyril Gavoille - University of Bordeaux
David Peleg - Weizmann Institute
Stéphane Pérennes - Nice Sophia-Antipolis
Ran Raz - Weizmann Institute

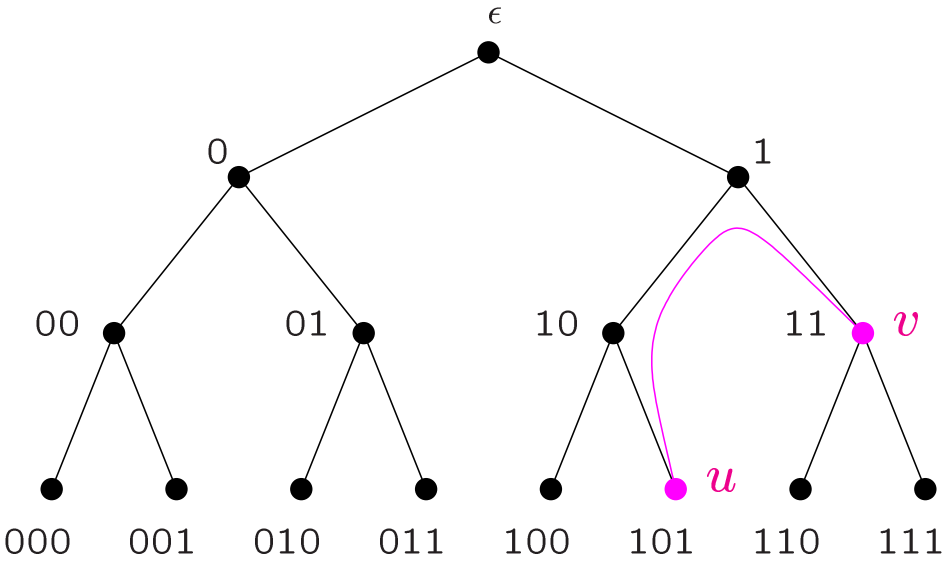
January 7, 2001

Problem:

How to label the nodes of a graph G in such a way that the distance between any two nodes u, v of G can be computed by inspecting the labels of u and v only?

(Applications to *Distributed Computing*, routing and broadcasting in large communication networks ...)

Ex: $\boldsymbol{G}$ = complete binary tree with $\boldsymbol{n = 2^{h+1} - 1}$ nodes



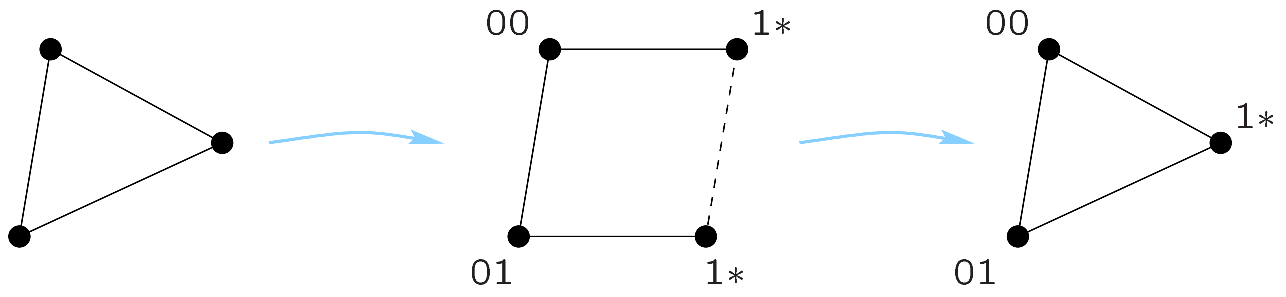
Labeling: node x of level i labeled with $L(x) \in \{0,1\}^i$, $i \leqslant h \leqslant \log n$

Distance decoder: $L(u) = \mathbf{1}01$, $L(v) = \mathbf{1}1$, $\text{prefix}(L(u), L(v)) = \mathbf{1}$

$$\begin{aligned} d_G(u,v) &= \text{length}(L(u)) + \text{length}(L(v)) - 2\text{length}(\text{prefix}(L(u),L(v))) \\ &= 3 + 2 - 2 * 1 = 3 \end{aligned}$$

General Case

Embed G in the "*Squashed Cube*"



$q = 2$

Labeling: $L(x) \in \{0, 1, *\}^q$

Distance: Hamming distance with $d(0, *) = d(1, *) = d(*, *) = 0$

Graham-Pollak '78: The *squashed-cube dimension of a connected graph G is at least r, where r is the maximum of the number of positive eigenvalues and number of negative eigenvalues of the distance matrix of G. (Conjecture: $q_{opt} \leqslant n - 1$.)*

Definition

Distance Labeling Scheme for $\mathcal{F}$ (family of connected graphs), is a pair $\langle L, f \rangle$ such that: $\forall G \in \mathcal{F}$, $\forall u \neq v \in V(G)$

(Labeling)	$L(u,G) \in \{1,\ldots,2^k\}$
(Decoder)	$f(L(u,G),L(v,G)) = d_G(u,v)$

Goals:

- Minimize k (the size in bits of the labels)
- Minimize distance query time (find a low worst-case time complexity decoder f, say constant time)

General Bounds

$O(n \log n)$ bit labels, $O(1)$ query time

	node	1	2	...	n
$L(u,G)$:	$\mathrm{dist}_G(u,\cdot)$	d_1	d_2	...	

$\longleftarrow$ $n \log n$ bits $\longrightarrow$

Actually, [Winkler'84], Graham-Pollack's Conjecture ($q_{opt} \leqslant n-1$)
$\Rightarrow$ $1.58n$ bit labels and $O(n)$ query time.

[In this paper]
$\Rightarrow$ $11n$ bit labels and $O(\log\log n)$ query time.

Lower Bounds

Let $\mathcal{F}_n$ be a family of n-node graphs (node labeled by $\{1,\ldots,n\}$), and $A \subseteq \{1,\ldots,n\}$. $\mathcal{F}_n$ is an A-family if $\forall\, G,H \in \mathcal{F}_n$, $\exists\, x,y \in A$ such that $\mathrm{dist}_G(x,y) \neq \mathrm{dist}_H(x,y)$. (Refinable).

Theorem. *Any $\langle L,f\rangle$ on an A-family $\mathcal{F}_n$ requires some labels of size*

$$\frac{1}{|A|} \cdot \log|\mathcal{F}_n| \ \textit{bits.}$$

Proof. Sequence of labels: $a_i \in A$

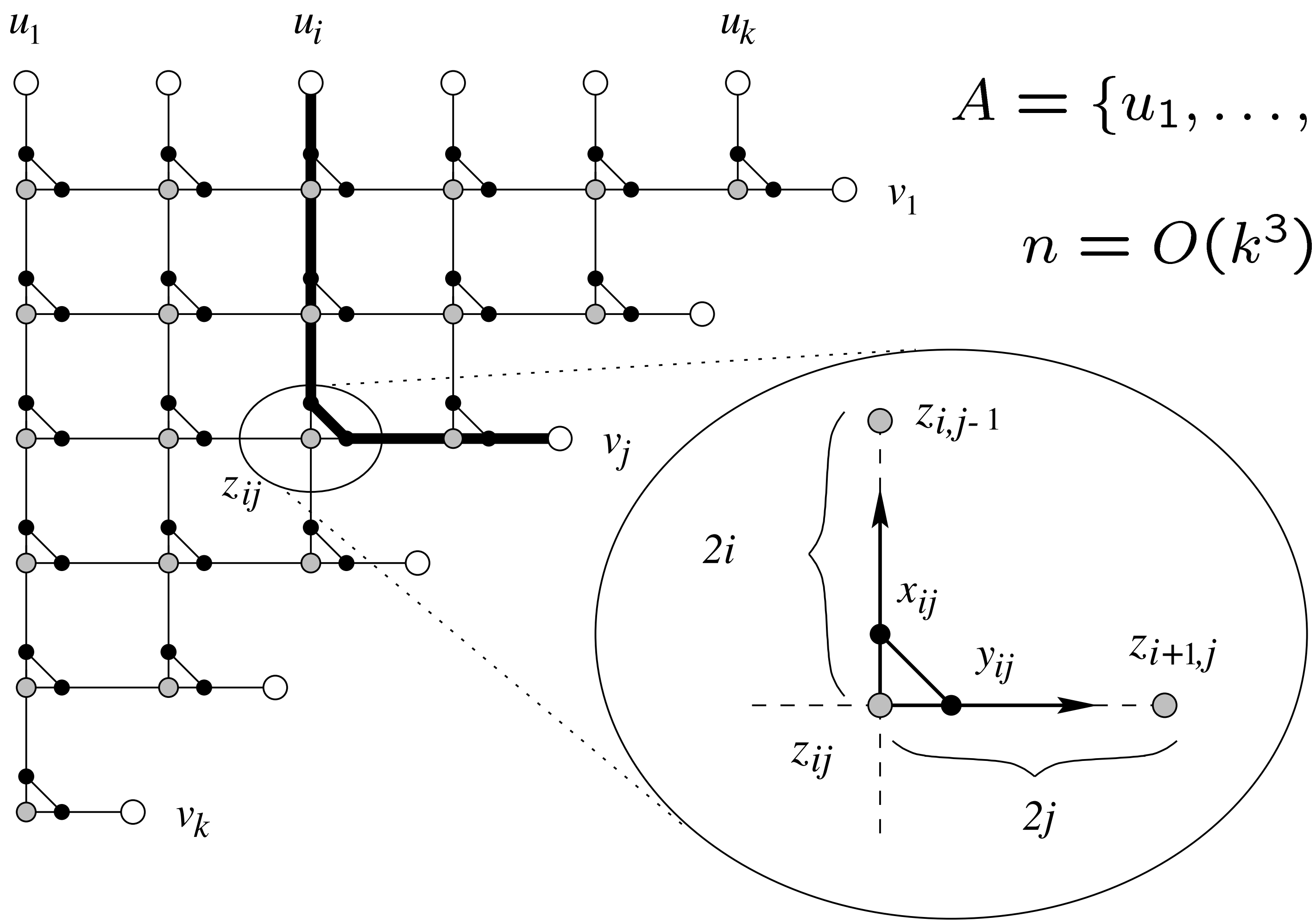
$$\begin{aligned} S(G) &= \langle L(a_1,G), L(a_2,G),\ldots)\rangle \\ S(H) &= \langle L(a_1,H), L(a_2,H),\ldots)\rangle \end{aligned}$$

$G \neq H \Rightarrow S(G) \neq S(H) \Rightarrow |\mathcal{F}_n|$ sequences
$\Rightarrow \exists a \in A$, $\exists G \in \mathcal{F}_n$, $|L(a,G)| \geqslant (\log|\mathcal{F}_n|)/|A|$. $\square$

Applications of the Lower Bound Theorem

- $\Omega(n)$ bit labels for general graphs ($A =$ set of all the nodes)
- $\Omega(\sqrt{n})$ bit labels for bounded degree graphs
- $\Omega(r(n))$ bit labels for graphs with $r(n)$-separator
- $\Omega(n^{1/3})$ bit labels for planar graphs ...

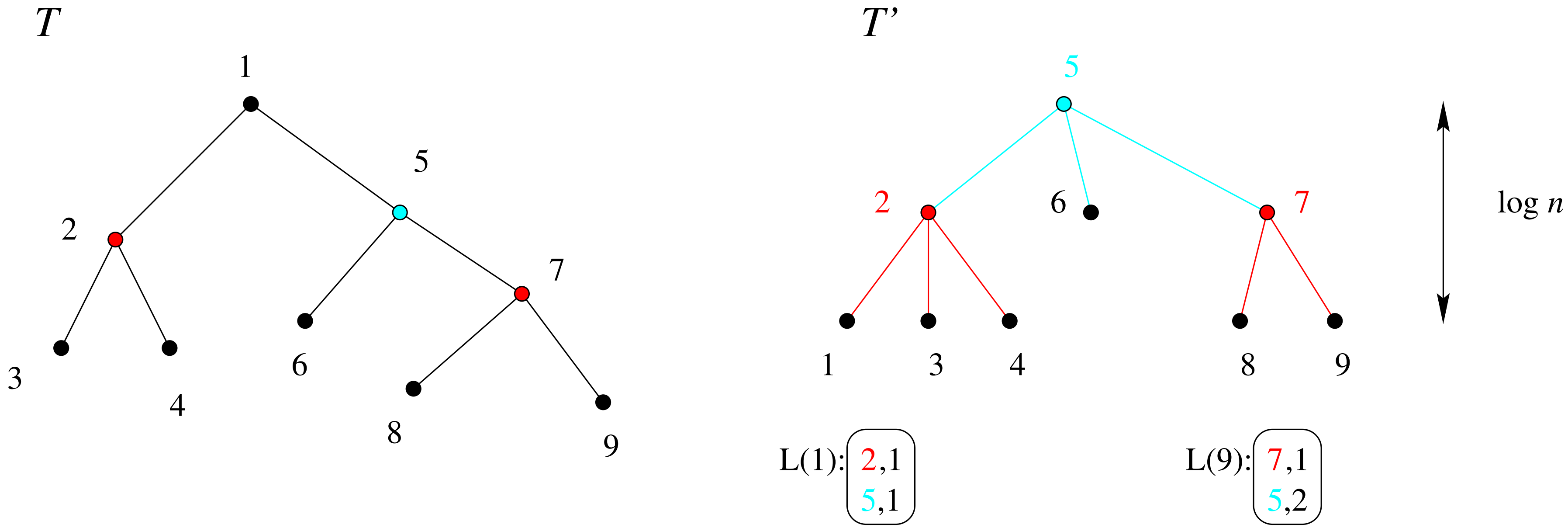
$$|\mathcal{F}_n| = 2^{\Theta(k^2)}$$



$$A = \{u_1, \ldots, u_k, v_1, \ldots, v_k\}$$

$$n = O(k^3),\ |A| = 2k$$

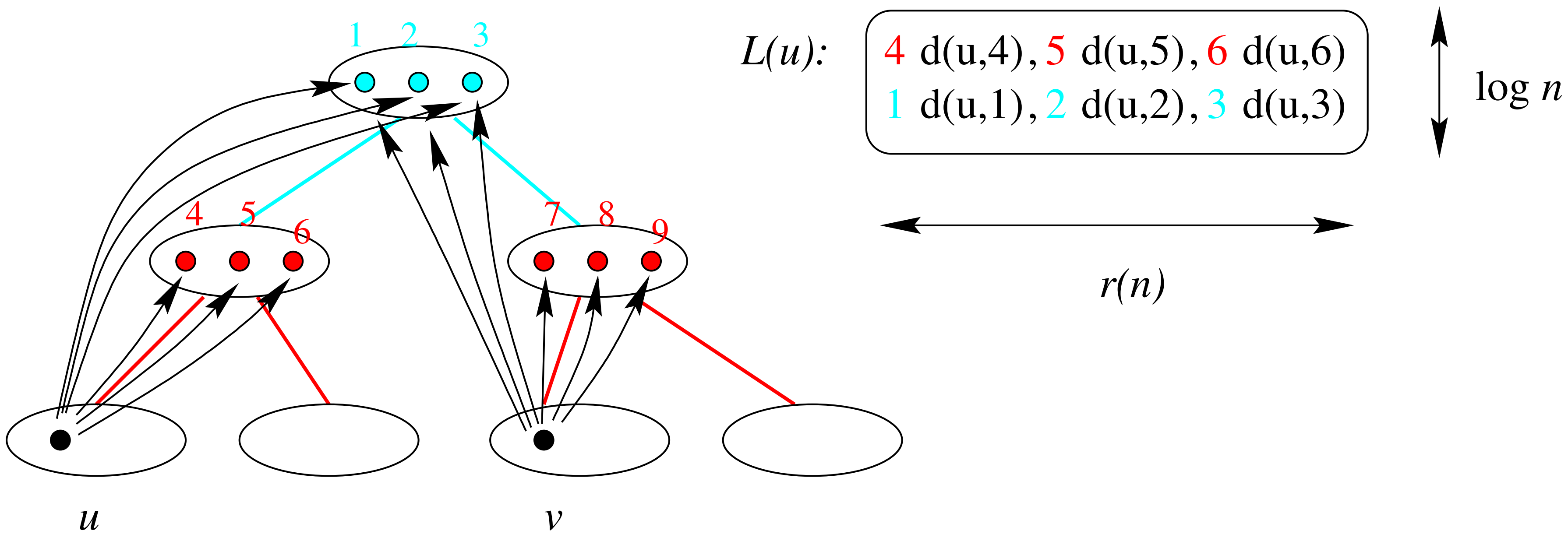
Upper Bound: Trees and Generalization



$2\log^2 n$ bit labels and $O(\log n)$ query time.

[Query time: $O(\log\log n)$, and $O(1)$ if allowing integer multiplication and division.]

Graphs with $r(n)$-separator



Graphs with $r(n)$-separator: $O(R(n)\log n)$ bit label
(where $R(n) := \sum_{i=0}^{\log n} r(n/2^i) \leqslant r(n)\log^2 n$)

$\Rightarrow$ bounded path-width: $O(\log^2 n)$ bit labels

$\Rightarrow$ bounded tree-width: $O(\log^2 n)$ bit labels

$\Rightarrow$ planar (and bounded genus): $O(\sqrt{n}\log n)$ bit labels

Lower Bound for Trees?

$\mathcal{T}_n$ = trees with node set $\{1, \ldots, n\}$

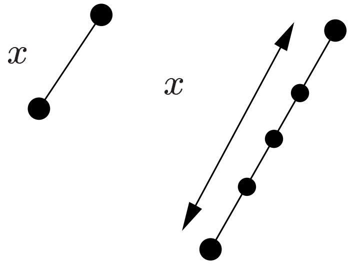
$|\mathcal{T}_n| = n^{n-2}$, A-family, $A = \{1, \ldots, n\}$

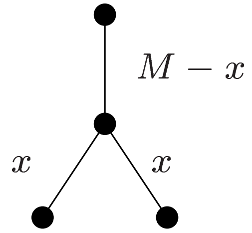
$$\Rightarrow \quad \frac{1}{n} \cdot \log |\mathcal{T}_n| \;=\; \log n - O(1)\ .$$

Lower Bound for Trees (cont'd)

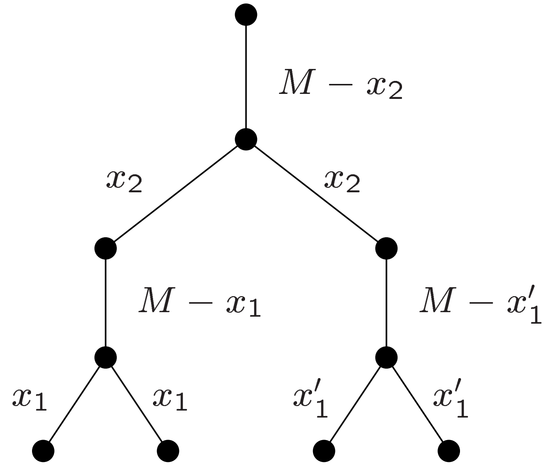
Family $\mathcal{T}_{h,M}$

$M \sim \sqrt{n}$
$h \sim \log n$





a $\mathcal{T}_{1,M}$ tree



#total of labels
for all binary trees
is at least $M^{h/2}$

$\mathcal{T}_{h,M}$

	1	2	3	...	n
T_1	ℓ_1	ℓ_2	ℓ_3	...	
T_2	ℓ_i	ℓ_1	ℓ_j	...	
T_3	ℓ_4	ℓ_i	ℓ_k	...	
$\vdots$		...	ℓ_t	...	

$\Rightarrow \exists\ \ell_t$ of size $\frac{h}{2} \log M \ \geqslant\ \frac{1}{8} \log^2 n$ bits.

Other Results

- Approximation of the distance

$$f(L(x,G),L(y,G)) \;\in\; [\mathrm{dist}_G(x,y),\, s\cdot\mathrm{dist}_G(x,y)]$$

  General graphs: $n/4$-lower bound for $s < 3$.

  Trees (and bounded treewidth): $\Theta(\log n \cdot \log\log n)$ (optimal) for $s = 1 + 1/\log n \to 1$.

- $\exists$ graphs with possibly short labels ($3\log n$), but such that if labels are shorter than $n/2$ bits, then the query time is necessarily greater than

$$2^{2^{\cdot^{\cdot^{\cdot^{2^n}}}}}$$

  (bounded resource Turing Machine)

Open Problems

- Label size for planar graphs?

$$\Omega(n^{1/3}) \ \ldots \ O(\sqrt{n}\log n)$$

- Label size for bounded degree graphs?

$$\Omega(\sqrt{n}) \ \ldots \ O(n)$$

- Label size for planar graphs for $s = O(1)$ approximation?

$$\Omega(?) \ \ldots \ O(n^{1/3}\log n) \text{ with } s = 3$$