Demand-Aware Network Designs of Bounded Degree

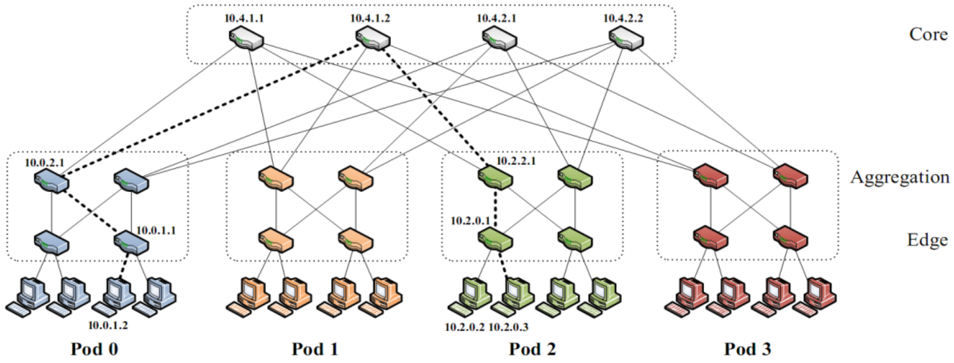
Chen Avin, Kaushik Mondal and Stefan Schmid

German-Israeli Foundation for Scientific Research and Development

Ben Gurion University of the Negev

Motivation

- Traditional datacentre networks are static:
  - Either over provisioned or under provisioned
  - Example. Fat-tree topologies: provide full bisection bandwidth

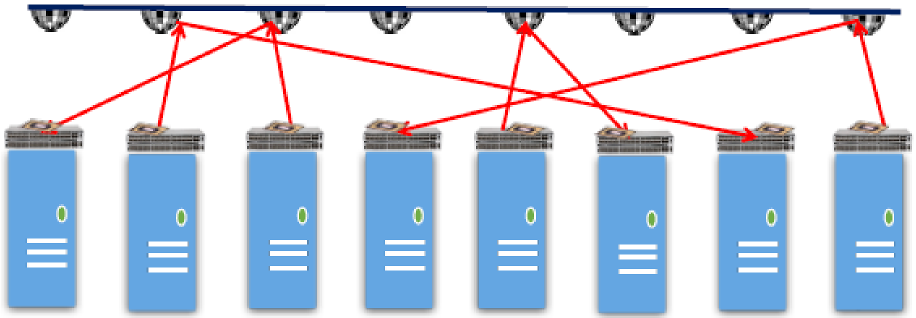


Al-Fares et al. "A scalable, commodity data center network architecture," ACM SIGCOMM Computer Communication Review, 2008

Motivation

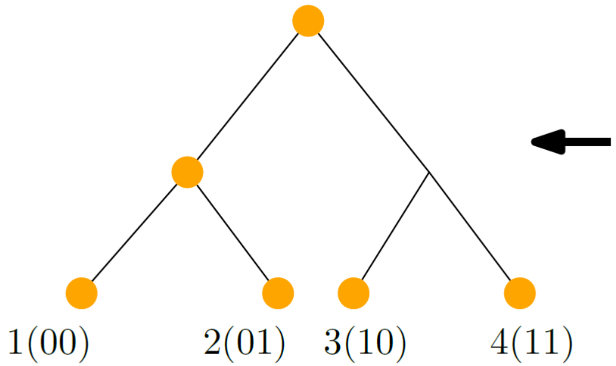
- New: Reconfigurable network
  - Reconfigurable on demand
  - Example: ProjecTor

M. Ghobadi et al. Projector: Agile reconfigurable data center interconnect. In Proc. ACM SIGCOMM, 2016



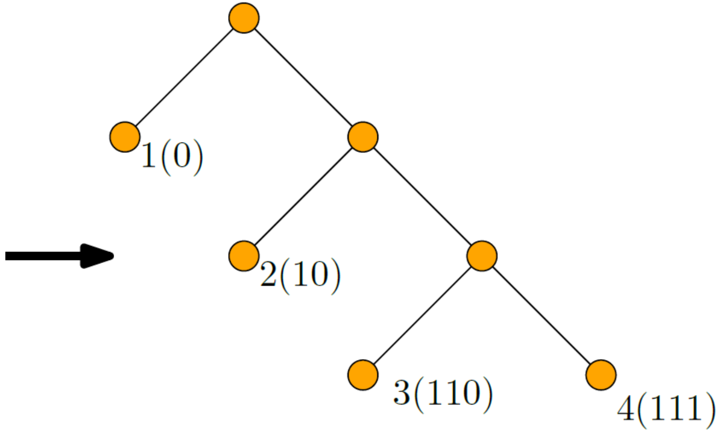
Motivation

- Instead of optimizing *worst-case* performance, design better networks with more information on communication patterns.
  - Example: Huffman coding.



$p(x_i)$	x_i
1/4	1
1/4	2
1/4	3
1/4	4

$p(x_i)$	x_i
1/2	1
1/4	2
1/8	3
1/8	4

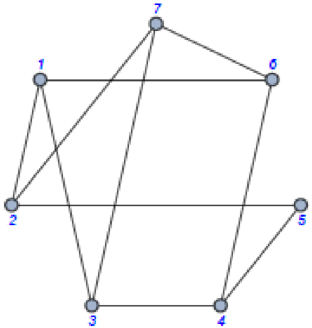


The Problem

$\mathcal{D}$: distribution matrix

$$
\begin{array}{c|ccccccc|}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
\hline
1 & 0 & \frac{2}{65} & \frac{1}{13} & \frac{1}{65} & \frac{1}{65} & \frac{2}{65} & \frac{3}{65} \\
2 & \frac{2}{65} & 0 & \frac{1}{65} & 0 & 0 & 0 & \frac{2}{65} \\
3 & \frac{1}{13} & \frac{1}{65} & 0 & \frac{2}{65} & 0 & 0 & \frac{1}{13} \\
4 & \frac{1}{65} & 0 & \frac{2}{65} & 0 & \frac{4}{65} & 0 & 0 \\
5 & \frac{1}{65} & 0 & \frac{3}{65} & \frac{4}{65} & 0 & 0 & 0 \\
6 & \frac{2}{65} & 0 & 0 & 0 & 0 & 0 & \frac{3}{65} \\
7 & \frac{3}{65} & \frac{2}{65} & \frac{1}{13} & 0 & 0 & \frac{3}{65} & 0 \\
\hline
\end{array}
$$

N: Demand aware network (DAN)



Bounded degree $\Delta=3$

Expected Path Length: $\mathrm{EPL}(\mathcal{D},\mathrm{N}) = \sum_{(u,v)\in\mathcal{D}} p(u,v)\cdot d_N(u,v)$

Bounded Network Design (BND)

- **Inputs**: Communication distribution $\mathcal{D}[\mathrm{p(i,j)}]_{\mathrm{nxn}}$ and a maximum degree Δ.

- **Output**: A Demand Aware Network $\mathrm{N} \in \mathcal{N}_\Delta$ s.t.

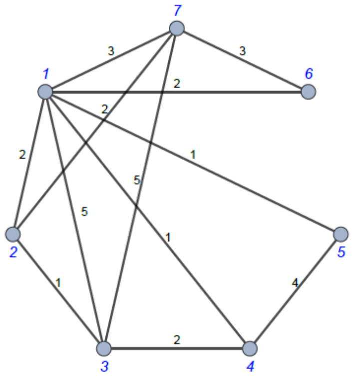
$$\mathrm{BND}(\mathcal{D}, \Delta) = \min_{\mathrm{N} \in \mathcal{N}_\Delta} \mathrm{EPL}(\mathcal{D}, \mathrm{N})$$

Related Problems: Embedding

- Embedding problem (guest graph-host graph)
  - $\Delta = 2$: Minimum linear arrangement problem
  - $\Delta > 2$: Type of host graph is not fixed beforehand
    - Flexible! Easier or harder?

$$
\begin{array}{c}
\begin{array}{cccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7
\end{array}\\
\begin{array}{c} 1\\2\\3\\4\\5\\6\\7 \end{array}
\begin{bmatrix}
0 & \frac{2}{65} & \frac{1}{13} & \frac{1}{65} & \frac{1}{65} & \frac{2}{65} & \frac{3}{65}\\
\frac{2}{65} & 0 & \frac{1}{65} & 0 & 0 & 0 & \frac{2}{65}\\
\frac{1}{13} & \frac{1}{65} & 0 & \frac{2}{65} & 0 & 0 & \frac{1}{13}\\
\frac{1}{65} & 0 & \frac{2}{65} & 0 & \frac{4}{65} & 0 & 0\\
\frac{1}{65} & 0 & \frac{3}{65} & \frac{4}{65} & 0 & 0 & 0\\
\frac{2}{65} & 0 & 0 & 0 & 0 & 0 & \frac{3}{65}\\
\frac{3}{65} & \frac{2}{65} & \frac{1}{13} & 0 & 0 & \frac{3}{65} & 0
\end{bmatrix}
\end{array}
$$

$\mathcal{D}$



$G_{\mathcal{D}}$

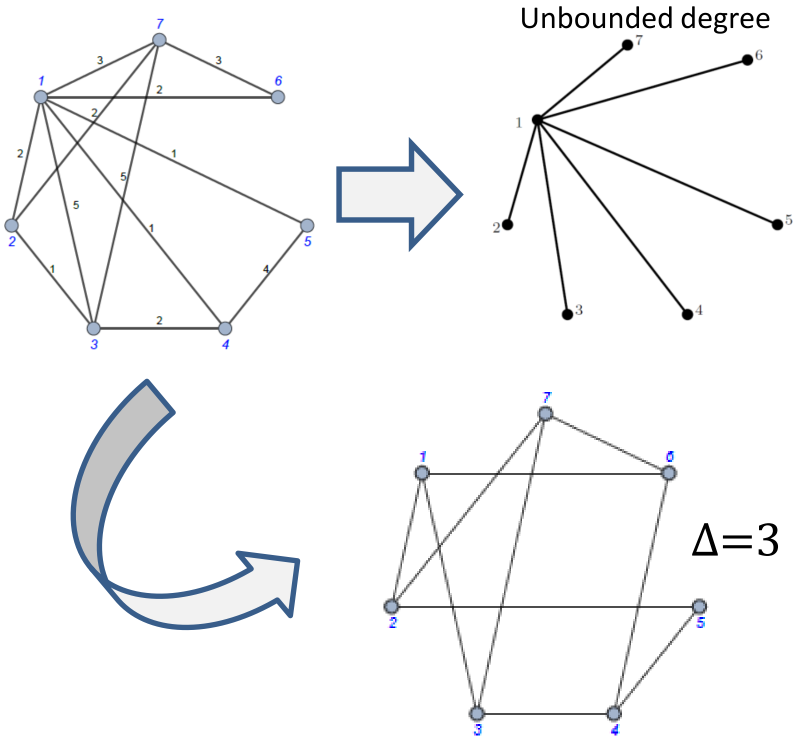
?

N

Related Problems: Spanners

- Spanners
  - Maintains local distortion
  - Presence of auxiliary edges like geometric spanner
  - Bounded degree

- Relation between spanner, entropy and BND!



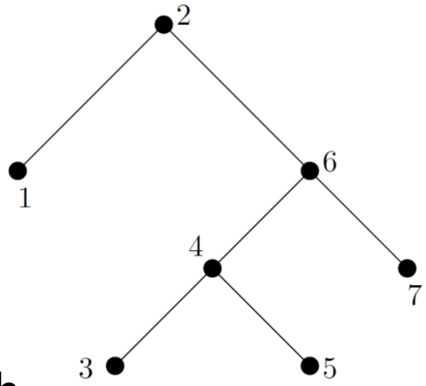
Related Problems: Coding

- Entropy and information theory
  - EPL(BST): bounded by entropy of destination frequencies $\bar{p}$

$$EPL(\bar{p},T) = \theta(H(X))$$

Kurt Mehlhorn. Nearly optimal binary search trees. Acta Inf., 1975.

  - Entropy: $H(X) = \sum_{i=1}^{n} p(x_i)\log_2(1/p(xi))$



- Coding theory
  - Huffman coding: Optimize expected code length which is equivalent to optimizing EPL in a BST.

Remainder of the talk

- Motivation
- The problem
- Related problems
- Lower bounds
- Bounded degree network designs
  - Tree distributions
  - Sparse distributions
  - Uniform and regular distributions
    - Locally doubling dimension distributions
- Contributions
- Future work

Lower bound

- **Theorem:** Let X, Y are distributed as marginal distribution of the sources and destinations in $\mathcal{D}$ respectively. Then

$$\mathrm{BND}(\mathcal{D}, \Delta) \geq \Omega(H_\Delta(Y|X) + H_\Delta(X|Y))$$

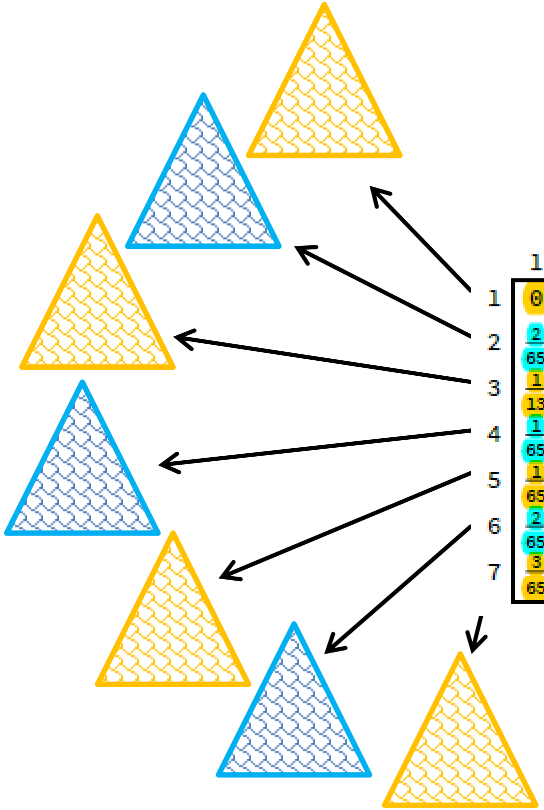
Marginal distribution:

	1	2	3	4	5	6	7	X
1	0	$\frac{2}{65}$	$\frac{1}{13}$	$\frac{1}{65}$	$\frac{1}{65}$	$\frac{2}{65}$	$\frac{3}{65}$	
2	$\frac{2}{65}$	0	$\frac{1}{65}$	0	0	0	$\frac{2}{65}$	
3	$\frac{1}{13}$	$\frac{1}{65}$	0	$\frac{2}{65}$	0	0	$\frac{1}{13}$	1/5
4	$\frac{1}{65}$	0	$\frac{2}{65}$	0	$\frac{4}{65}$	0	0	
5	$\frac{1}{65}$	0	$\frac{3}{65}$	$\frac{4}{65}$	0	0	0	
6	$\frac{2}{65}$	0	0	0	0	0	$\frac{3}{65}$	
7	$\frac{3}{65}$	$\frac{2}{65}$	$\frac{1}{13}$	0	0	$\frac{3}{65}$	0	
Y					1/13			

- $H(X|Y) = \sum_{i=1}^{n} p(x_i, y_j) \log_2(1/p(x_i|y_j))$

Lower bound

- **Proof idea** (EPL=$\Omega(H_\Delta(Y|X))$):
- Build optimal Δ-ary tree for each source i.
- Consider union of all trees.
- Violates degree restriction but valid lower bound.

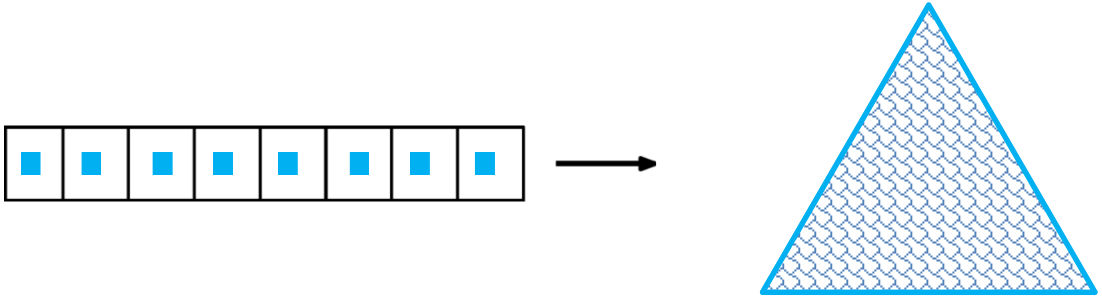


	1	2	3	4	5	6	7
1	0	$\frac{2}{65}$	$\frac{1}{13}$	$\frac{1}{65}$	$\frac{1}{65}$	$\frac{2}{65}$	$\frac{3}{65}$
2	$\frac{2}{65}$	0	$\frac{1}{65}$	0	0	0	$\frac{2}{65}$
3	$\frac{1}{13}$	$\frac{1}{65}$	0	$\frac{2}{65}$	0	0	$\frac{1}{13}$
4	$\frac{1}{65}$	0	$\frac{2}{65}$	0	$\frac{4}{65}$	0	0
5	$\frac{1}{65}$	0	$\frac{3}{65}$	$\frac{4}{65}$	0	0	0
6	$\frac{2}{65}$	0	0	0	0	0	$\frac{3}{65}$
7	$\frac{3}{65}$	$\frac{2}{65}$	$\frac{1}{13}$	0	0	$\frac{3}{65}$	0

Lower bound

- Helping **Lemma**: Let T^i_Δ be an optimal $\Delta-$ary tree built for the normalized distribution of i’th row $\mathcal{D}[i]$ of $\mathcal{D}$. Then

$$\mathrm{EPL}(\mathcal{D}[i], T^i_\Delta) \geq \Omega(H_\Delta((Y \mid X=i)))$$



$\mathcal{D}[i]$ T^i_Δ

Lower bound

- Considering union of n trees,

$$\mathrm{EPL}(\mathcal{D},\mathrm{N}_{\Delta}) \geq \sum_{i=1}^{n} \mathrm{p}(\mathrm{i})\, \mathrm{EPL}(\mathcal{D}[\mathrm{i}], \mathrm{Ti}_{\Delta})$$

$$\geq \textstyle\sum_{i=1}^{n} \mathrm{p}(\mathrm{i})\, (\mathrm{H}_{\Delta}(\mathrm{Y} \,|\mathrm{X}{=}\mathrm{i}))$$

$$= \Omega(\mathrm{H}_{\Delta}(\mathrm{Y} \,|\mathrm{X}))$$

Lower bound

- Similarly, for incoming communications

$$\mathrm{EPL}(\mathcal{D},\mathrm{N}_\Delta) \geq \Omega(\mathrm{H}_\Delta(\mathrm{X}\,|\mathrm{Y}))$$

- Hence, $\mathrm{BND}(\mathcal{D},\Delta) \geq \Omega(\mathrm{H}_\Delta(\mathrm{Y}|\mathrm{X}) + \mathrm{H}_\Delta(\mathrm{X}|\mathrm{Y}))$

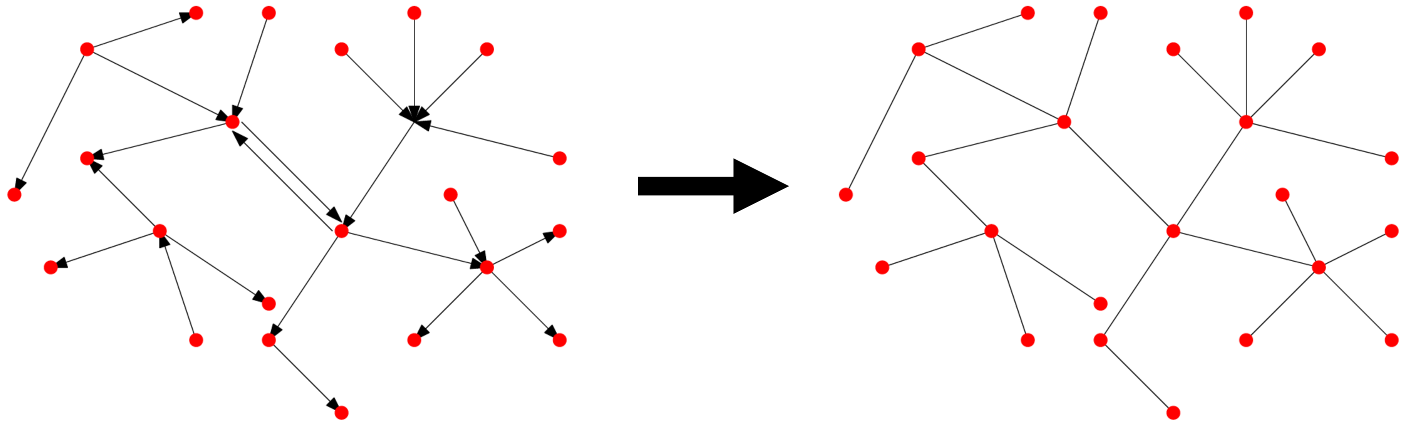
Remainder of the talk

- Motivation
- The problem
- Related problems
- Lower bounds
- Bounded degree network designs
  - Tree distributions
  - Sparse distributions
  - Uniform and regular distributions
    - Locally doubling dimension distributions
- Contributions
- Future work

Tree distributions

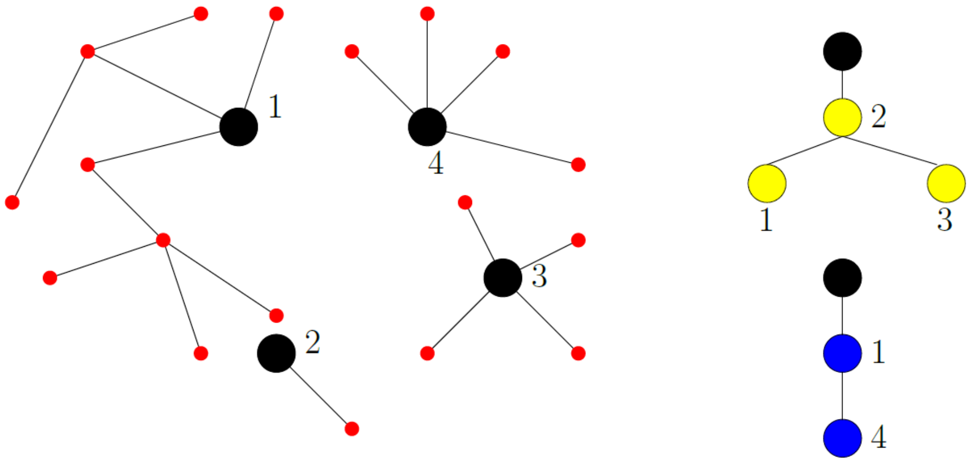
- **Theorem:** Let $\mathcal{D}$ be such that $G_{\mathcal{D}}$ is a tree (ignoring the edge direction). It is possible to generate a DAN N with maximum degree 8, such that,

$$\mathrm{EPL}(\mathcal{D},\mathrm{N}) \leq \mathrm{O}(\mathrm{H}(\mathrm{Y}|\mathrm{X})+\mathrm{H}(\mathrm{X}|\mathrm{Y}))$$



Tree distributions

- **Proof idea:**
  - Parent-child relationship implied by arbitrary root.
  - Arrange the outgoing children of each node in a binary tree.
  - Repeat it for the incoming edges.
  - Degree as parent: at most 2
  - Degree as child: at most 6



Tree distributions

- $\mathrm{EPL}(\mathcal{D},\mathrm{N}) \leq \sum_{i=1}^{n} p(i)\,\mathrm{EPL}(\overrightarrow{c_i},\overrightarrow{B_i}) + \sum_{i=1}^{n} q(i)\,\mathrm{EPL}(\overleftarrow{c_i},\overleftarrow{B_i})$

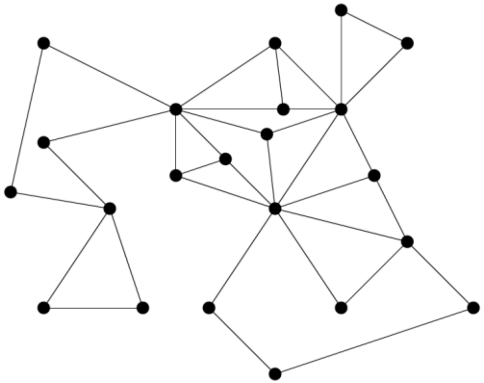
$$\leq \sum_{i=1}^{n} p(i)\,\mathrm{H}(\overrightarrow{D_i}) + \sum_{i=1}^{n} q(i)\,\mathrm{H}(\overleftarrow{D_i})$$

$$= \mathrm{H}(Y|X)+\mathrm{H}(X|Y)$$

- $\overrightarrow{D_i}$: normalized destination distributions of v_i (i'th row)
- Helping **Lemma:** Let $\bar{p}$ be the destination distributions for a source node then it is possible to find a Δ-ary tree T, s.t.,

$$\mathrm{EPL}(\bar{p},\mathrm{T}) \leq O(\mathrm{H}_{\Delta}(\bar{p}))$$

Sparse distributions



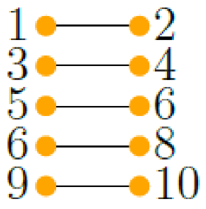
- Real datacentre's traffic shows evidence that the demand distributions are indeed sparse.

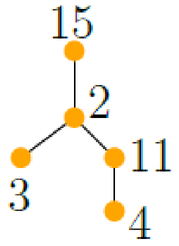
- **Theorem:** $G_{\mathcal{D}}$ is a sparse graph with average degree Δ_{avg} , then it is possible to find a DAN N with maximum degree $12\Delta_{avg}$, such that

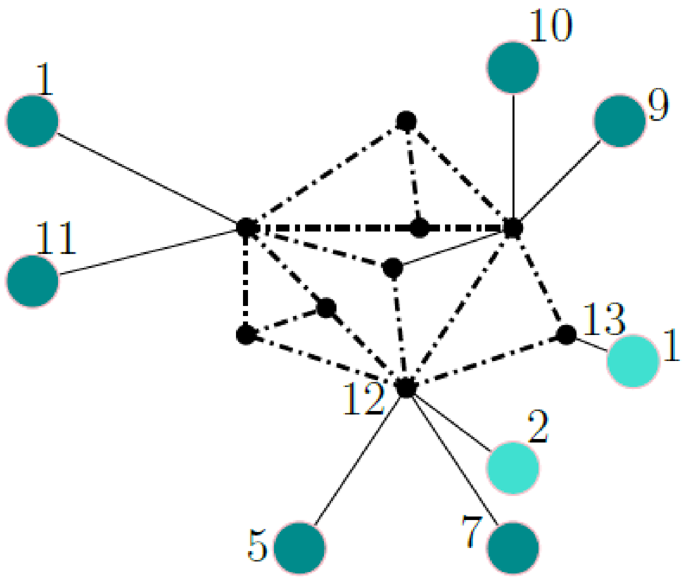
$$EPL(\mathcal{D},N) \leq O(H(Y|X)+H(X|Y))$$

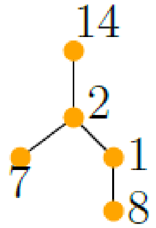
Sparse distributions

- Find low degree nodes.
- Mark low-low edges.
- High degree nodes with all low degree neighbors.
- Make binary tree of them.
- Low degree node between high-high edge.
- High nodes have only low neighbours. Make tree.









Sparse distributions

- **Proof:** Node degree<12 Δ_{avg}
  - n/2 low degree nodes with degree bounded by $2\Delta_{\text{avg}}$.
    - Each low degree node in $2\Delta_{\text{avg}}$ trees so degree at most $6\Delta_{\text{avg}}$ in N.
  - Each low degree node helps at most Δ_{avg} edges
    - so present in another 2 Δ_{avg} trees and hence degree increases by another $6\Delta_{\text{avg}}$

Sparse distributions

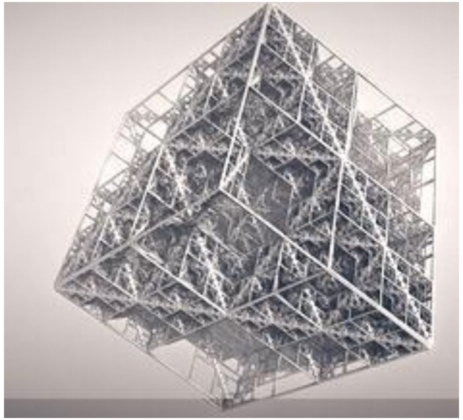
- Create new matrix $\mathcal{D}'$

- $\mathrm{EPL}(\overrightarrow{\mathcal{D}'[\mathrm{i}]}, \overrightarrow{\mathrm{B}_\mathrm{i}}) \leq \mathrm{O}(\mathrm{H}(\mathrm{Y}|\mathrm{X}=\mathrm{i}))$ and
$\mathrm{EPL}(\overleftarrow{\mathcal{D}'[\mathrm{j}]}, \overleftarrow{\mathrm{B}_\mathrm{j}}) \leq \mathrm{O}(\mathrm{H}(\mathrm{X}|\mathrm{Y}=\mathrm{j}))$

- Remaining analysis is almost similar to tree distribution.

Regular and uniform distributions

- Regular graph can be dense
  - Example: Hypercube



- **Theorem:** $\mathcal{D}$ is uniform, regular and possibly dense. If $G_{\mathcal{D}}$ has a constant sparse (graph) spanner, then $\exists$ DAN N such that,

$$\mathrm{EPL}(\mathcal{D},\mathrm{N}) \leq \mathrm{O}(\mathrm{H}(\mathrm{Y}|\mathrm{X})+\mathrm{H}(\mathrm{X}|\mathrm{Y}))$$

Regular and uniform distributions

- **Proof:** Maximum degree of spanner S is r, if $\mathcal{D}$ is r-regular.

$$\mathrm{EPL}(\mathcal{D},N) = \sum_{(u,v)\in \mathcal{D}} p(u,v) d_N(u,v)$$
$$\leq \sum_{(u,v)\in \mathcal{D}} p(u,v) d_S(u,v)\ 2 \log r$$
$$= \mathrm{EPL}(\mathcal{D},S) \cdot 2 \cdot \log r = O(\log r) = O(H(Y \mid X))$$

- **Lemma**: If S has average degree Δ_{avg} and maximum degree Δ_{max}, then $\exists$ S' with maximum degree $8\Delta_{avg}$ such that

$$d_{s'}(u,v) \leq 2 \log \Delta_{max}\, d_s(u,v)$$

Regular and uniform distributions

- **Corollary.** $\mathcal{D}$: constant and regular communication distribution. Possible to generate $\mathrm{DAN}\ \ \mathrm{N}$ if,
  - If $\mathrm{G}_{\mathcal{D}}$ is a hypercube with $\mathrm{n} \log \mathrm{n}$ edges
    - has sparse 3-spanner
  - If $\mathrm{G}_{\mathcal{D}}$ is a (possibly dense) chordal graph
    - has constant sparse spanner

Regular and uniform distributions

- A special case: Possible to generate DAN N, if $G_{\mathcal{D}}$ has minimum degree $\Delta \geq n^{1/c}$, for any constant c.

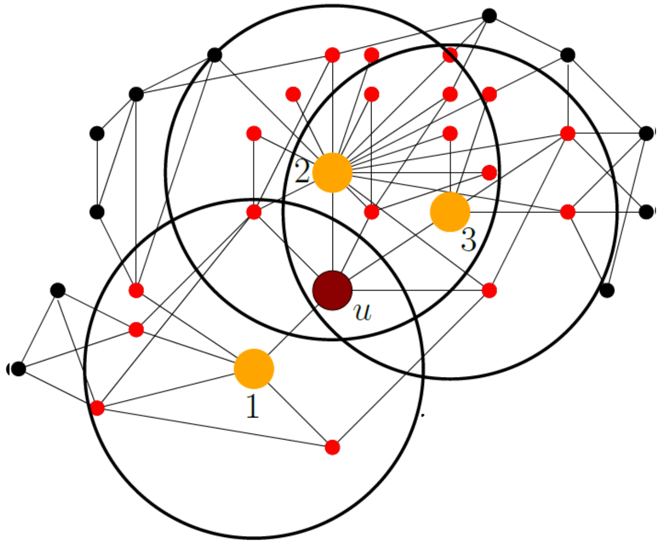
  - Create a Δ-ary tree with the nodes of $G_{\mathcal{D}}$ and call it N
  - Distortion $\log_{\Delta} n$ on N
  - $\mathrm{EPL}(\mathcal{D}, N) = \sum_{(u,v) \in \mathcal{D}} p(u,v) d_N(u,v)$

$$\leq \sum_{(u,v) \in \mathcal{D}} p(u,v) d_G(u,v) \cdot 2 \log n = O(H_{\Delta}(Y \mid X))$$

Since, $H_{\Delta}(Y \mid X) \geq (1/c) \log_{\Delta} n$

Locally-bounded doubling dimension

- **LDD**: $G_{\mathscr{D}}$ has a Locally-bounded Doubling Dimension (LDD) iff 2-hop neighbours are covered by 1-hop neighbours of finite nodes.
- Formally, $B(u, 2) \subseteq \cup_{i=1}^{\lambda} B(yi, 1)$
- LDD vs BDD:
  - Every BDD is a LDD.
  - Dense, unbounded degree, irregular.



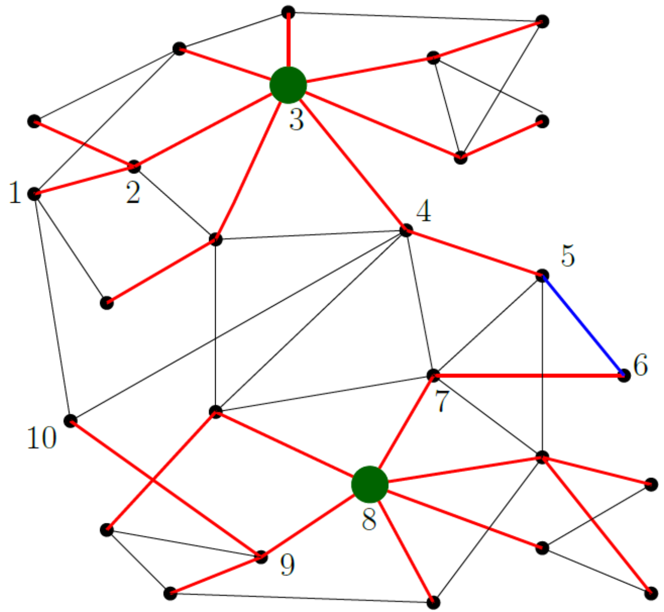
Locally-bounded doubling dimension

- **Lemma:** There exists a sparse 9-spanner for LDD. This is also a subgraph spanner.

- **Def.(ε-net):** A subset V' of V is a **ε**-net for a graph $G = (V,E)$ if
  - for every $u, v \in V'$, $d_G(u, v) > \boldsymbol{\varepsilon}$
  - for each $w \in V$, $\exists$ at least one $u \in V'$ such that, $d_G(u,w) \leq \boldsymbol{\varepsilon}$

Locally-bounded doubling dimension

- **Proof idea:** Find a 2-net and add nodes to one of the closest 2-net nodes.
- Join two clusters if there are edges in between.
- Distortion 9
- Sparse: Only finite number of net nodes within 5 hops.



Locally-bounded doubling dimension

- **Theorem:** It is possible to find a $\mathrm{DAN}\ \mathrm{N}$ for an uniform and regular locally doubly dimension graph such that,

$$\mathrm{EPL}(\mathcal{D},\mathrm{N}) \leq \mathrm{O}(\mathrm{H}(\mathrm{Y}|\mathrm{X})+\mathrm{H}(\mathrm{X}|\mathrm{Y}))$$

- **Proof:** Existence of constant sparse spanner.

Contributions

- BND is a fundamental problem
- Provide a lower bound in terms of entropy
- Matching upper bound for sparse distribution, uniform and regular distributions.
- Convert network to low degree network s.t. $\mathrm{EPL} \leq O(H(Y|X))$

Future work

- More general graphs: regular/maximum degree $\mathrm{n}^{1/\mathrm{r}}$, for any r.
- Do we require alternate flavours of graph entropy?
- Maintaining the bounded degree network dynamically.

Thank You !