Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit

J. A. Tropp and A. C. Gilbert

IEEE Trans. on Inform. Theory

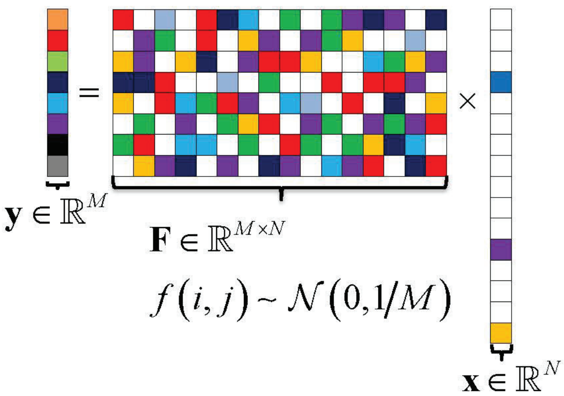
Presenter : Sangjun Park

GIST, Dept. of Information and Communication, INFONET Lab.

Gwangju Institute of Science and Technology

Questions and System Model

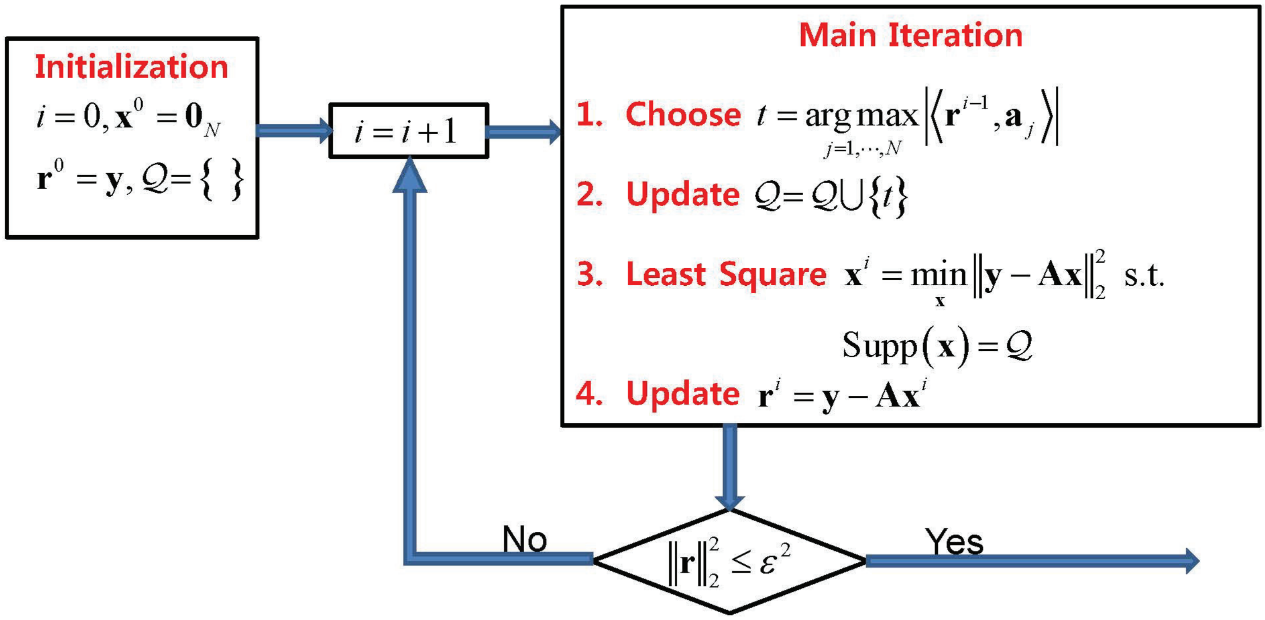
- Let us suppose that we aim to find the support set of a sparse vector by using OMP.
- Then, what is a sufficient condition for successful OMP?



Orthogonal Matching Pursuit

- OMP finds one index at a time for approximating the solution of

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ subject to } \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2 \le \varepsilon^2$$



Sufficient conditions for successful OMP

- There are many papers that report sufficient conditions for successful OMP.

Year	A sufficient condition	Types
2004	$\mu < 1/(2K-1)$	Deterministic
2010	$\delta_{K+1} < 1/(3\sqrt{K})$	Deterministic
2012	$\delta_{K+1} < 1/(\sqrt{K}+1)$	Deterministic
This paper	$M=\Omega(K\log(N))$	Probabilistic

2007: J. Tropp, "Greed is good: Algorithmic results for sparse approximation," IEEE Trans. On. Inform. Theory
2010: M. A. Davenport, M. B. Wakin, "Analysis of Orthogonal Matching Pursuit Using the Restricted Isometry Property", IEEE Trans. On. Inform. Theory
2012: J. Wang and B. Shim, "On the recovery Limit of Sparse Signals Using Orthogonal Matching Pursuit", IEEE Trans. Signal Processing Letter

The short overview of the paper [2012]

- To derive their sufficient condition, the authors considered the event that OMP correctly selects index j at the i^{th} iteration.
- The event occurs if $\min_{t\in\mathcal{I}}\left\|\langle \mathbf{a}_t,\mathbf{y}\rangle\right\|_2 > \max_{t\notin\mathcal{I}}\left\|\langle \mathbf{a}_t,\mathbf{y}\rangle\right\|_2$.
- They have shown that the left term is lower bounded by

$$\min_{t\in\mathcal{I}}\left\|\langle \mathbf{a}_t,\mathbf{y}\rangle\right\|_2 \geq \frac{1}{\sqrt{K}}\left(1-\delta_K\right)\left\|\mathbf{x}_{\mathcal{I}}\right\|_2.$$

- Also, they have shown that the right term is upper bounded by

$$\max_{t\notin\mathcal{I}}\left\|\langle \mathbf{a}_t,\mathbf{y}\rangle\right\|_2 \leq \left(1-\delta_{K+1}\right)\left\|\mathbf{x}_{\mathcal{I}}\right\|_2.$$

- Then, they have derived their sufficient condition from the two bounds.

The main Theorem

- (OMP with Admissible Measurement matrix.) Fix $\delta \in (0,1)$, and choose $M = \Omega(Klog(N/\delta))$. Suppose that $\mathbf{x}$ is an arbitrary $K-$sparse vector in $\mathcal{R}^N$, and draw a random $M \times N$ admissible measurement matrix $\mathbf{A}$ independent from the vector. Given the measurement vector $\mathbf{y} = \mathbf{A}\mathbf{x}$. Then, OMP can reconstruct the support set with probability exceeding $1-\delta$.

Admissible Measurement Matrices

- An admissible measurement matrix for K −sparse vectors in $\mathcal{R}^N$ is an $M \times N$ random matrix $\mathbf{A}$ with four properties.

  (M0) Independence : The columns of $\mathbf{A}$ are stochastically independent.

  (M1) Normalization : $\mathbb{E}\left[\|\mathbf{a}_i\|_2^2\right] = 1$ for $j = 1, \cdots, N$.

  (M2) Joint correlation : Let $\{\mathbf{u}^t\}$ be a sequence of vectors whose l_2 norms do not exceed one. Let $\mathbf{a}$ be a column of $\mathbf{A}$ that is independent from $\{\mathbf{u}^t\}$. Then,

  $$\mathbb{P}\left\{\max_t \left|\left\langle \mathbf{a}, \mathbf{u}^t \right\rangle\right| \le \varepsilon\right\} \ge 1 - 2K\exp\left(-c\varepsilon^2 M\right)$$

  (M3) Smallest singular value: Given an $M \times K$ submatrix $\boldsymbol{Z}$ from $\mathbf{A}$, the largest singular value $\sigma_{min}(\mathbf{Z})$ satisfies $\mathbb{P}\left\{\sigma_{\min}(\mathbf{Z}) \ge 0.5\right\} \ge 1 - \exp(-cM)$

The proof of the main Theorem-1

- First, let us define the greedy ratio at the l^{th} iteration:

$$\rho\left(\mathbf{r}^l\right) := \frac{\max_{i \notin \mathcal{I}} \left|\left\langle \mathbf{r}^l, \mathbf{a}_i \right\rangle\right|}{\max_{i \in \mathcal{I}} \left|\left\langle \mathbf{r}^l, \mathbf{a}_i \right\rangle\right|}$$

- OMP correctly selects an index belonging to the support set if $\rho\left(\mathbf{r}^l\right) < 1$.

- OMP correctly reconstructs the support set when the event $E_{succ} := \max_{l \le K} \rho\left(\mathbf{r}^l\right) < 1$ occurs

- We aim to obtain the probability

$$\mathbb{P}\left\{E_{succ}\right\} := \mathbb{P}\left\{\max_{l \le K} \rho\left(\mathbf{r}^l\right) < 1\right\}$$

$$\ge \mathbb{P}\left\{\max_{l \le K} \rho\left(\mathbf{r}^l\right) < 1 \cap \sigma_{\min}\left(\mathbf{A}_{\mathcal{I}}\right) \ge 0.5\right\}$$

- Owing to (M3), we can solve LS within the K^{th} iterations.

The proof of the main Theorem-2

- Continuously, we aim to consider the probability

$$\mathbb{P}\left\{\max_{l\le K}\rho\left(\mathbf{r}^l\right)<1\middle|\sigma_{\min}\left(\mathbf{A}_{\mathcal{I}}\right)\ge 0.5\right\}$$

- For this end, we consider the greedy ratio at the l^{th} iteration. Then, we have

$$\rho\left(\mathbf{r}^l\right)=\frac{\max_{i\notin\mathcal{I}}\left|\left\langle\mathbf{r}^l,\mathbf{a}_i\right\rangle\right|}{\max_{i\in\mathcal{I}}\left|\left\langle\mathbf{r}^l,\mathbf{a}_i\right\rangle\right|}=\frac{\max_{i\notin\mathcal{I}}\left|\left\langle\mathbf{r}^l,\mathbf{a}_i\right\rangle\right|}{\left\|\mathbf{A}_{\mathcal{I}}^T\mathbf{r}^l\right\|_\infty}\le\frac{\sqrt{K}\max_{i\notin\mathcal{I}}\left|\left\langle\mathbf{r}^l,\mathbf{a}_i\right\rangle\right|}{\left\|\mathbf{A}_{\mathcal{I}}^T\mathbf{r}^l\right\|_2}$$

- Now, we simplify the upper bound of the greedy ratio. First, let us define $\mathbf{r}^l:=\mathbf{u}^l\left\|\mathbf{A}_{\mathcal{I}}^T\mathbf{r}^l\right\|_2/0.5$. Then, the upper bound becomes

$$\frac{\sqrt{K}\max_{i\notin\mathcal{I}}\left|\left\langle\mathbf{r}^l,\mathbf{a}_i\right\rangle\right|}{\left\|\mathbf{A}_{\mathcal{I}}^T\mathbf{r}^l\right\|_2}=\frac{\sqrt{K}\max_{i\notin\mathcal{I}}\left|\left\langle\mathbf{u}^l\left\|\mathbf{A}_{\mathcal{I}}^T\mathbf{r}^l\right\|_2/0.5,\mathbf{a}_i\right\rangle\right|}{\left\|\mathbf{A}_{\mathcal{I}}^T\mathbf{r}^l\right\|_2}$$
$$=2\sqrt{K}\max_{i\notin\mathcal{I}}\left|\left\langle\mathbf{u}^l,\mathbf{a}_i\right\rangle\right|.$$

The proof of the main Theorem-3

- Owing to M3, we have $\left\|\mathbf{A}_{\mathcal{I}}^T\mathbf{r}^l\right\|_2/\left\|\mathbf{r}^l\right\|_2\ge\sigma_{\min}\left(\mathbf{A}_{\mathcal{I}}\right)\ge 0.5.$
- Then, we can show that the l^2 norm of the vector $\mathbf{u}^l$ is always less than one.

$$\mathbf{u}^l=0.5\,\mathbf{r}^l/\left\|\mathbf{A}_{\mathcal{I}}^T\mathbf{r}^l\right\|_2\le\mathbf{r}^l/\left\|\mathbf{r}^l\right\|_2$$

- Now, we have

$$\mathbb{P}\left\{\max_{l\le K}\rho\left(\mathbf{r}^l\right)<1\middle|\sigma_{\min}\left(\mathbf{A}_{\mathcal{I}}\right)\ge 0.5\right\}\ge\mathbb{P}\left\{\max_{l\le K}2\sqrt{K}\max_{i\notin\mathcal{I}}\left|\left\langle\mathbf{u}^l,\mathbf{a}_i\right\rangle\right|<1\middle|\sigma_{\min}\left(\mathbf{A}_{\mathcal{I}}\right)\ge 0.5\right\}$$
$$=\mathbb{P}\left\{\max_{i\notin\mathcal{I}}\max_{l\le K}\left|\left\langle\mathbf{u}^l,\mathbf{a}_i\right\rangle\right|<\frac{1}{2\sqrt{K}}\middle|\sigma_{\min}\left(\mathbf{A}_{\mathcal{I}}\right)\ge 0.5\right\}$$
$$\ge\prod_{i\notin\mathcal{I}}\mathbb{P}\left\{\max_{l\le K}\left|\left\langle\mathbf{u}^l,\mathbf{a}_i\right\rangle\right|<\frac{1}{2\sqrt{K}}\middle|\sigma_{\min}\left(\mathbf{A}_{\mathcal{I}}\right)\ge 0.5\right\}$$
$$\ge\left[1-2K\exp\left(-cM/(4K)\right)\right]^{N-K}$$

The proof of the main Theorem-4

- In addition, we have $\mathbb{P}\{\sigma_{\min}(\mathbf{Z}) \geq 0.5\} \geq 1-\exp(-cM)$.
- Thus, we finally obtain

$$\mathbb{P}\{E_{succ}\} \geq \left[1-2K\exp\left(-cM/(4K)\right)\right]^{N-K}\left[1-\exp(-cM)\right].$$

- To simplify the lower bound, we apply the inequality $(1-x)^n \geq 1-kn$ for $n \geq 1$ and $x \leq 1$. Then, for $K(N-K) \leq N^2/4$, we have

$$\mathbb{P}\{E_{succ}\} \geq 1-2K(N-K)\exp\left(-cM/(4K)\right)-\exp(-cM).$$

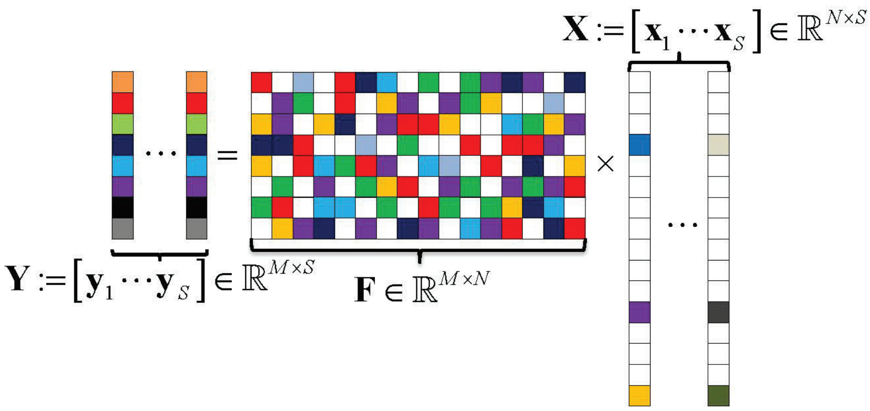
- By again simplifying the above lower bound, we have

$$\mathbb{P}\{E_{succ}\} \geq 1-N^2\exp(-cM/K).$$

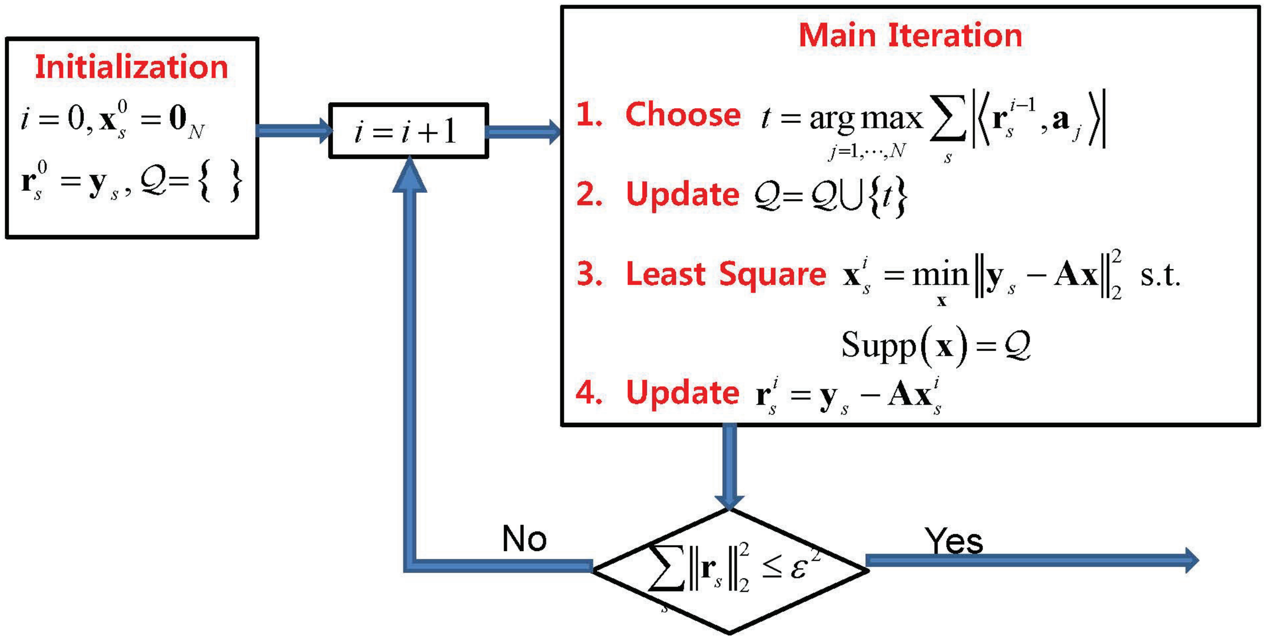
- Finally, we can see that the choice $M = \Omega(Klog(N/\delta))$ is sufficient to reduce the failure probability below δ.

New researches problems

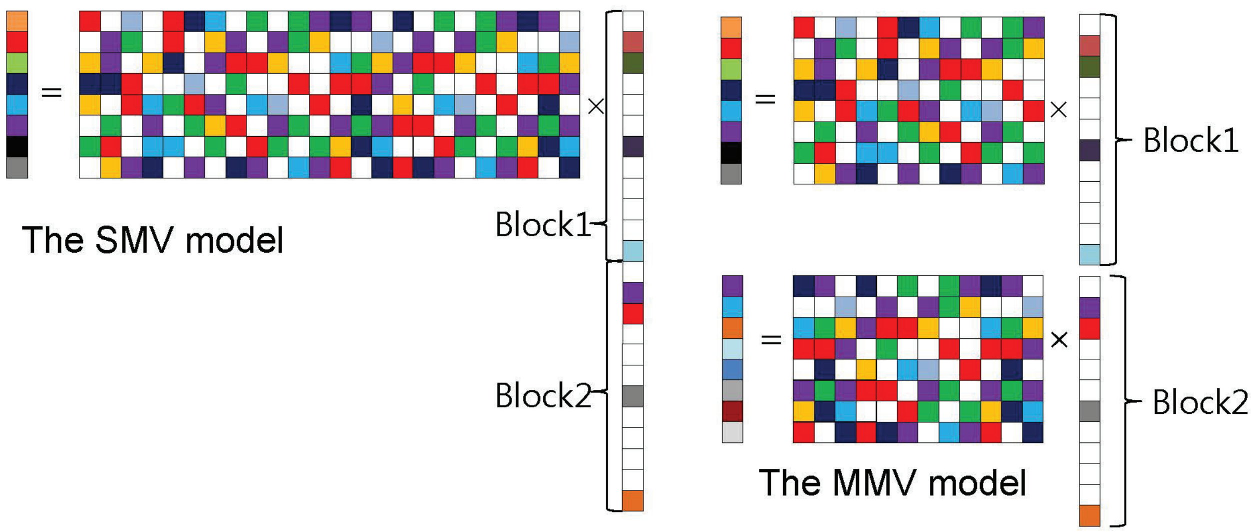
1. Can we establish a sufficient condition for Simultaneously Orthogonal Matching Pursuit?



Simultaneously Orthogonal Matching Pursuit



New researches problems



2. Let M_1 be the number of measurements in the SMV model when OMP is exploited. Let M_2 be the total number of measurements in the MMV model when SOMP is exploited. What is the relation between M_1 and M_2?

Link Status Monitoring Using Network Coding

M. H. Firooz et al.

To appear IEEE/ACM Trans. on Networking

Presenter : Jin-Taek Seong

GIST, Dept. of Information and Communications, INFONET Lab.

Gwangju Institute of Science and Technology

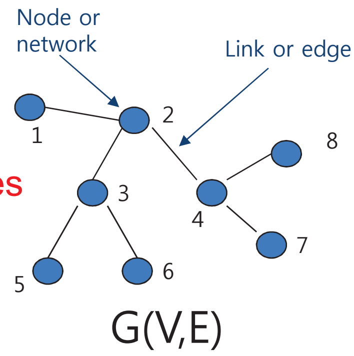
Journal Club Meeting, June 27, 2013

Outline

- **Network Tomography**
  - Introduction (Network Monitoring)
  - Approaches:
    - Deterministic vs. Stochastic
    - Active vs Passive
  - Challenges: Overhead, Identifiability
- **Network Coding**
  - Applications to network monitoring: new method
  - Optimization : speed/complexity tradeoffs

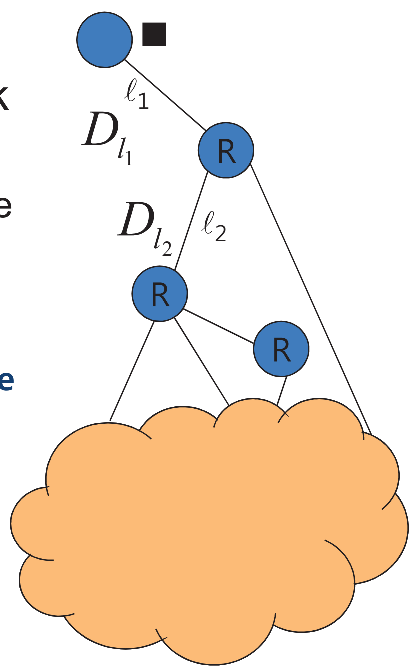
Network Tomography

- Networks: **set of nodes, links modeled as graph G(V,E)**
- Network monitoring
  - Involves collection of network performance statistics (link delay, link loss or failure status)
  - Important for QoS guarantees (media streaming, interactive video applications)
- Challenges
  - Choice of appropriate measurement techniques and algorithms



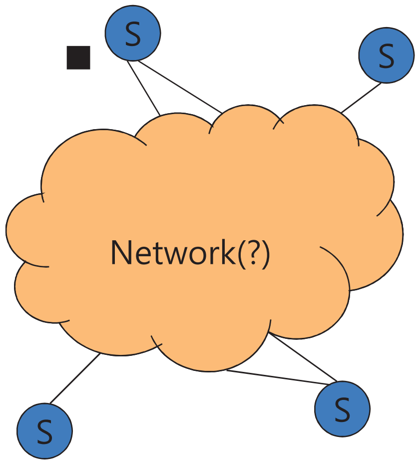
Measurement Methods

- **Node-oriented**: These methods are based on cooperation among network nodes, e.g., ping or traceroute
  - Using Ping, round trip delay to every node can be measured.
  - Uses Internet control message protocol (ICMP) packets
    - **Many routers do NOT respond to these packets**
  - Many service providers do not own the entire network

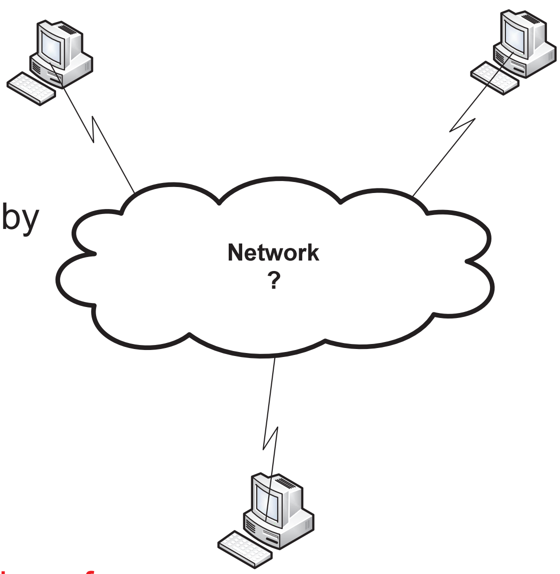


Measurement Methods

- **Edge-oriented**: Access is available to all nodes at the edge only (and not to any in the interior)
  - Does not require exchanging special control messages between interior nodes
  - ***Inverse problem***: estimate *link level* status from end-to-end (path level) measurements



Measurement Methods

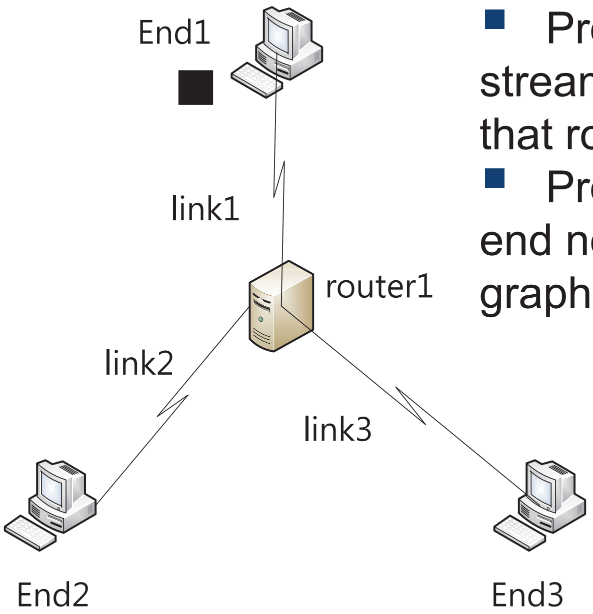


❖ **Active** (sending probe packets)
 - Adds overhead to normal data traffic by introducing new control packets

❖ **Passive** (insitu traffic analysis)
 - No overhead; temporal and spatial dependence might bias measurement

❖ Considered method: *edge-oriented, active* network tomography
  - Given a network, and a limited number of end hosts, **when** can we infer failure status of the links?

End-to-End Probing



- Probes are inserted into a data stream, and end-to-end properties on that route measured.
- Probes are exchanged between end nodes using routing matrix of the graph

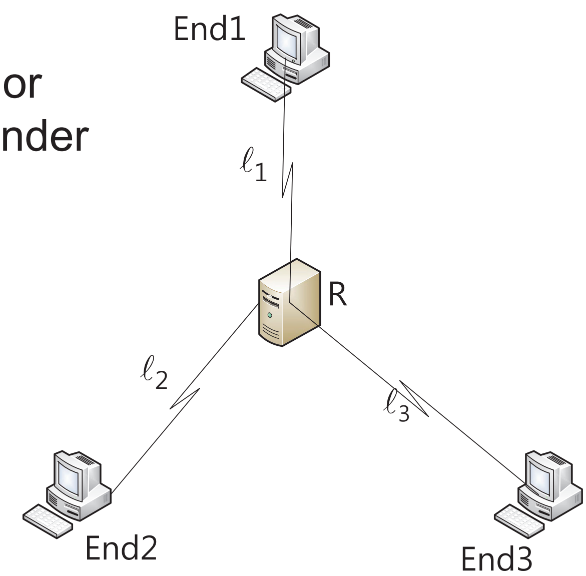
Routing matrix A

	$link1$	$link2$	$link3$
$End1 \to End2$	1	1	0
$End1 \to End3$	1	0	1
$End2 \to End3$	0	1	1

End-to-End Probes

- Routing matrix relates link attribute to route attribute
- For some parameters like delay or path loss, this relation is linear under some assumptions

$$\begin{bmatrix} D_{End1\to End2} \\ D_{End1\to End3} \\ D_{End2\to End3} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} D_{l_1} \\ D_{l_2} \\ D_{l_3} \end{bmatrix}$$



Deterministic

- Link attributes (e.g. delay) are considered unknown, constant
- Goal: estimate constants
- Link attributes are typically time varying

→ method is suitable for periods of local '*stationarity*'

Journal Club Meeting, June 27, 2013

Stochastic

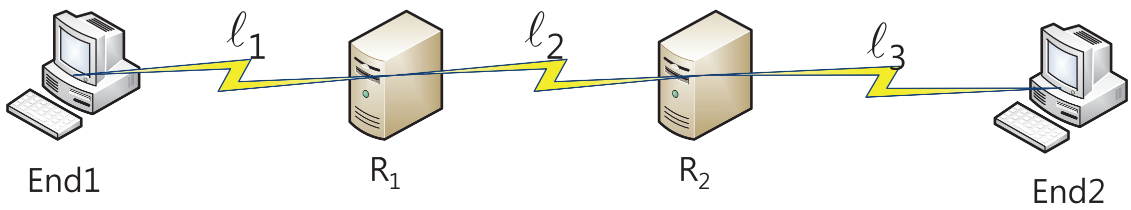
- Link attribute specified by a suitable probability distribution
  - e.g. link delay follows a Gaussian distribution
- Estimation problem: unknown model parameters

based on path observation in the presence of additive noise

Deterministic vs. Stochastic Methods

- Stochastic
  - Bayesian - requires a prior distribution
    - incorrect choice leads to biases in the estimates
  - More computationally intensive
- Deterministic
  - Lower complexity but suffers from generic *identifiability* (will be discussed later) problems

Link Failure Model

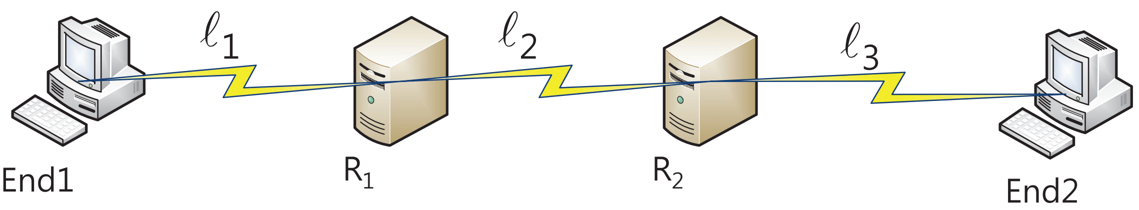


Define an indicator function for status of each link

$$x_{l_i} = \begin{cases} 0 & l_i \text{ is ok} \\ 1 & l_i \text{ is congested} \end{cases}$$

$$y_{end1 \to end2} = \begin{cases} 0 & \text{all of } l_1, l_2, l_3 \text{ is ok} \\ 1 & o.w. \end{cases}$$

Binary Deterministic Model



$$y_{end1 \rightarrow end2} = x_{l_1} \ or \ x_{l_2} \ or \ x_{l_3}$$

$$\mathrm{y = Ax}$$

A: N-by-M binary routing matrix

x: M-by-1 binary vector, the status of each link

y: N-by-1 binary vector, the status of each path (measurements)

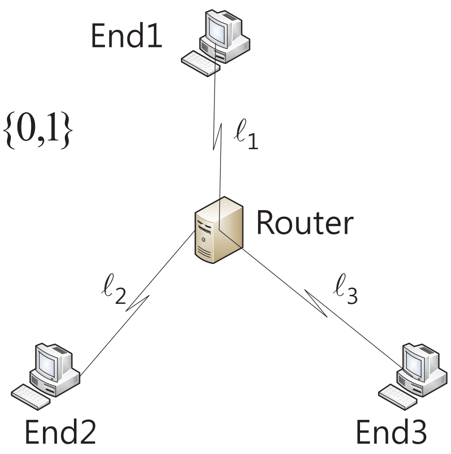
Failure Monitoring

- Network G(V,E) with set of paths *P*
- **x, y** are binary vectors
- A path is congested if at least one of its links is congested

$$\mathbf{x} \in \{0,1\}^{|\mathbf{E}|}, \mathbf{y} \in \{0,1\}^{|\mathbf{P}|}$$

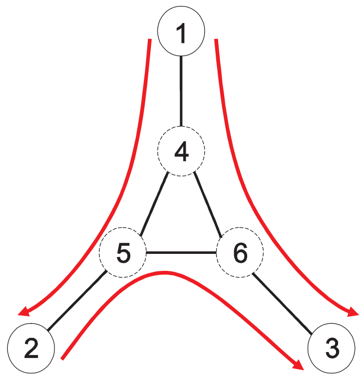
$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{matrix} End1 \rightarrow End2 \\ End1 \rightarrow End3 \\ End2 \rightarrow End3 \end{matrix} \overset{\begin{matrix} l_1 & l_2 & l_3 \end{matrix}}{\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}} \begin{bmatrix} x_{l_1} \\ x_{l_2} \\ x_{l_3} \end{bmatrix}, \quad x_{l_1} \in \{0,1\}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_{l_1}(OR)x_{l_2} \\ x_{l_1}(OR)x_{l_3} \\ x_{l_2}(OR)x_{l_3} \end{bmatrix}$$



Identifiability y = Ax

- **Problem:** Estimate **x** from **y** with
  - **A *(N-by-M)* :** binary routing matrix
  - **x *(M-by-1)* :** binary link failure status
  - **y *(N-by-1)* :** end-to-end measurements

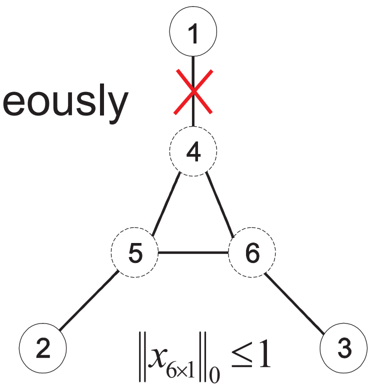


6 links, 3 End-to-End routes ➔ M=6, N=3

- **Identifiability:** a network is identifiable if **y = Ax** has a ***unique solution***
  - Usually, *M* (# of links in network) >> *N* (# of measurements), so network is generically NOT identifiable.

Identifiability: Binary Model

- Solution: limit (maximum) number of failed links inside the network
  - Suppose at most *k* links can fail simultaneously



$\|x_{6\times1}\|_0 \leq 1$

Only one link can be congested

- Definition: *k-Identifiability*
  - Network is k-identifiable if

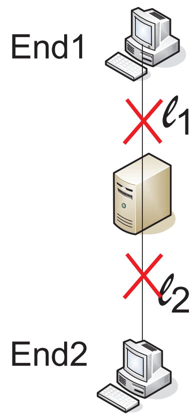
$$\|\mathbf{x}_{|E|\times1}\|_0 \leq k$$

$$\forall \mathbf{x}_1, \mathbf{x}_2 \text{ s.t.} \|\mathbf{x}_1\|_0 \leq k, \|\mathbf{x}_2\|_0 \leq k,\ \mathbf{x}_1 \neq \mathbf{x}_2 => \mathbf{A}\mathbf{x}_1 \neq \mathbf{A}\mathbf{x}_2$$

- From end-to-end observation it is possible to uniquely identify up to ***k*** congested links

1-Identifiability

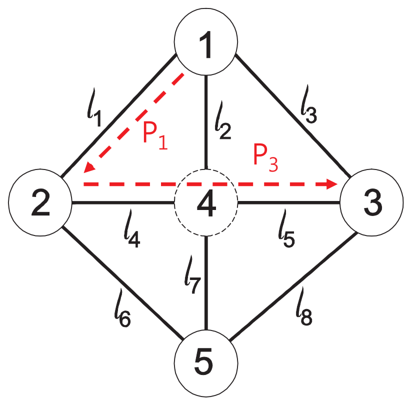
- ❖ *A network with an intermediate degree two node is **not** 1-identifiable*
  - ✓ If path End1→End2 is congested, it is impossible to determine which link among l_1 and l_2 is congested .
- Necessary but not sufficient!



$$x_{l_1} = 1 \Rightarrow y_{End1 \to End2} = 1$$
$$x_{l_2} = 1 \Rightarrow y_{End1 \to End2} = 1$$

k=1 Identifiability

- **1-identifiability Theorem**: End-to-End probe based measurements can detect a unique congested link in a network *if and only if* there are ***no two identical columns*** in the network routing matrix



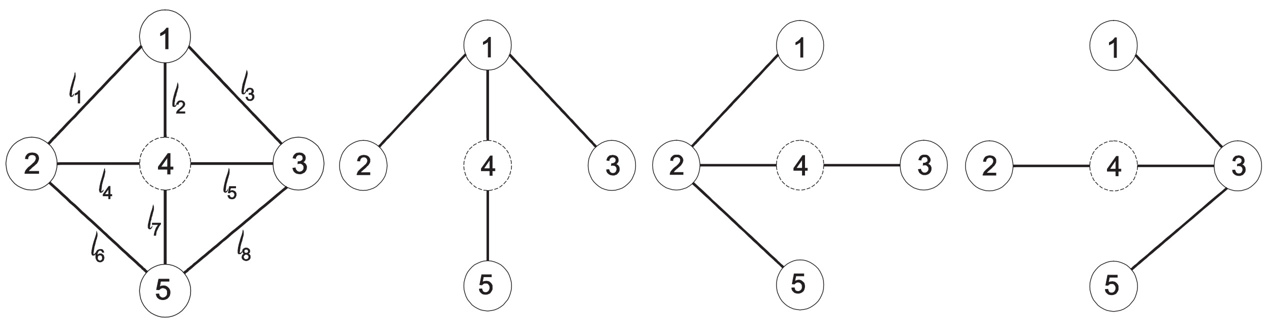
$$\begin{matrix} P_1 \\ \\ \\ P_3 \\ \\ \\ \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

k- identifiability

- **k-identifiability Theorem**: End-to-End probe based measurements can detect a unique congested link in a network *only if* there are ***no k+1 dependent columns*** in the network routing matrix

Shortest Path Routing Revisited

- Packets are sent on shortest path between two end nodes
  - sub-graphs = tree starting from a boundary (source) node
    - ➢ Node 4 has two degrees in all graphs
- But node 4 has 4 degrees in the original network



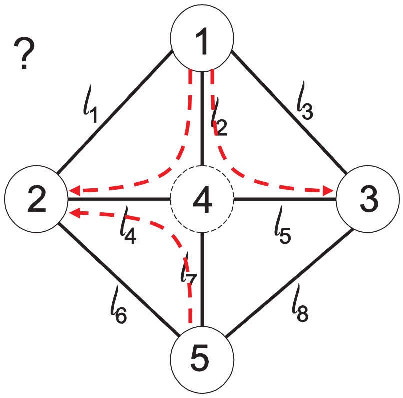
Revisiting Shortest Path Routing

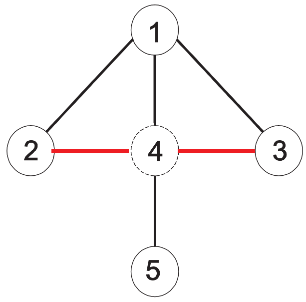
- What if we could change routing matrix ?

Example: in place of shortest path routing, route packets through longer paths, e.g. $n_1 \rightarrow \ell_2 \rightarrow \ell_4 \rightarrow n_2$

  – Now network is 1-identifiable !

- Intrinsic limitation for end-to-end measurement methods based on shortest path routes
  – probes transmitted along such paths contain only *minimum information*





Journal Club Meeting, June 27, 2013

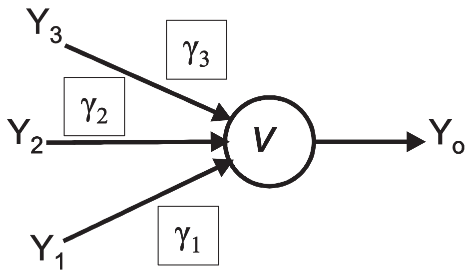
Solution

- Look to exchange probes between boundary nodes via other (non-shortest) paths?
- Changing the routing tables violates tomography assumption
- Use **Network Coding**; exploit broadcast nature of network coding, a transmitted probe will traverse almost every path between two boundary nodes

Linear Network Coding

- Network Coding is a coding at layer three
- The coding is conducted over the finite field F_u, $u=2^q$
- Each coded symbol can be represented by *q*-bits within an IP layer frame
- Signal $\boldsymbol{Y}$(j) on an outgoing link *j* of node *v* is a linear combination of signals $\boldsymbol{Y}$(i) on incoming link *i* of *v*:
  - We assume there is no process generated at node *v*

$$Y(j) = \sum_{\{l:d(l)=v\}} \gamma_l Y(l)$$



Received Symbols

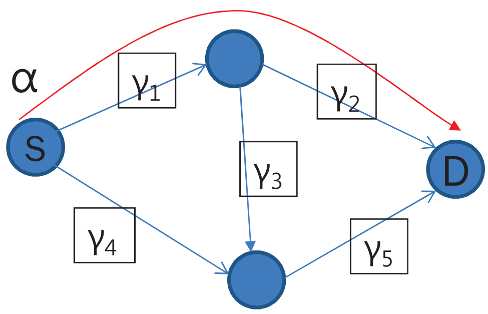
- $\boldsymbol{P}^i$: *i*-th route from source to destination
- Source sends α over $\boldsymbol{P}^i$

$$y = \alpha \prod_{l \in P^i} \gamma_l = \alpha \beta_i(G), \quad \alpha \in F_{2^q}$$

$$\beta_i(G) = \prod_{l \in P^i} \gamma_l \quad \text{Path NC Coef.}$$

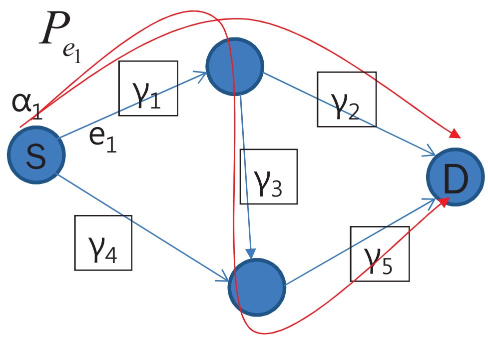
- β_i depends on topology *G* hence $\beta_i(G)$

$$y = \alpha \gamma_1 \gamma_2 = \alpha \beta_1(G)$$



Received Symbols: Linear Model

- e_k one of source outgoing links
- P_{ek} : collection of all paths between source and destination starts at the k-th outgoing edge e_k
- Source sends α_k over e_k. By superposition destination receives



$$y = \alpha_k \sum_{P^i \in P_{e_k}} \prod_{l \in P^i} \gamma_l = \alpha_k \sum_{i=1}^{|P_{e_k}|} \beta_{i,e_k}(G)$$

$$y = \alpha_1(\gamma_1\gamma_2 + \gamma_1\gamma_3\gamma_5) = \alpha_1(\beta_{1,e_1} + \beta_{2,e_1})$$

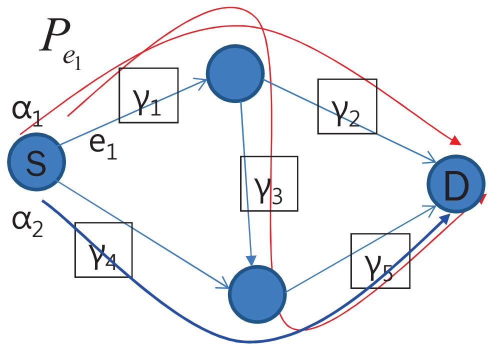
Journal Club Meeting, June 27, 2013

Received Symbols: Linear Model

- Source sends out symbols α_k over e_k using superposition once more

$$y = \sum_{k=1}^{K} \alpha_k \sum_{i=1}^{|P_{e_k}|} \beta_{i,e_k}(G)$$

- In vector format: y=**α**[t]**β(G)**
- **β(G)** is total network coding vector



$$y = \alpha_1(\gamma_1\gamma_2 + \gamma_1\gamma_3\gamma_5) + \alpha_2\gamma_4\gamma_5$$

Received Symbols: Linear Model

- Source sends symbols in *M* consecutive time slots:

$$y_{M\times 1} = A_{M\times N}\beta(G)_{N\times 1}$$

$$\beta(G)_{N\times 1} = \Big[\underbrace{\beta_{1,e_1}\ \ \beta_{2,e_1}\ \cdots\ \beta_{N_1,e_1}}_{P_{e_1}}\ \Big|\ \underbrace{\beta_{1,e_1}\ \cdots\ \beta_{N_2,e_2}}_{P_{e_2}}\ \Big|\ \cdots\ \Big|\ \underbrace{\beta_{N_K,e_K}}_{P_{e_K}}\Big]^t$$

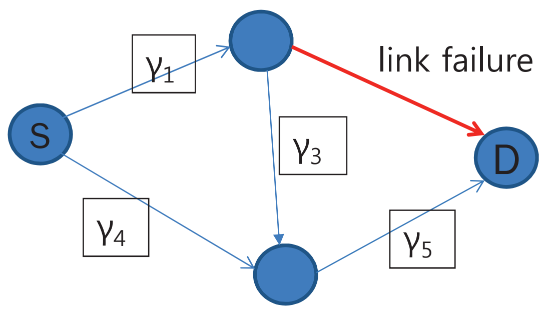
$$N = \left|\bigcup_{i=1}^{K}\mathcal{P}_{e_i}\right|$$

$$A_{M\times N} = \begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,|\mathcal{P}_{e1}|} & \cdots & \alpha_{1,N} \\ \alpha_{2,1} & \cdots & \alpha_{2,|\mathcal{P}_{e1}|} & \cdots & \alpha_{2,N} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \alpha_{M,1} & \cdots & \alpha_{M,|\mathcal{P}_{e1}|} & \cdots & \alpha_{M,N} \end{bmatrix}$$

A: consisting *K* distinct columns

Link Failure Model

- If a link is severely congested, packets are significantly delayed and assumed lost at the destination
- We model the network with link *l* in congestion state by its *edge deleted subgraph* denoted by $G_l(V,E_l)$



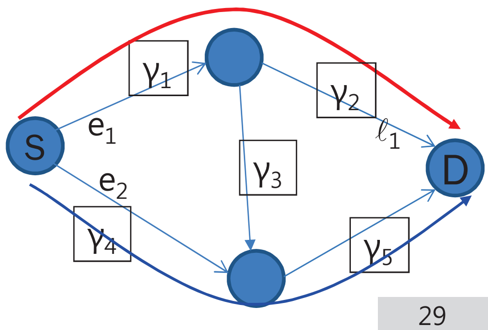
Link Failure Model

- Total network coding vector of $G_l(V;E_l)$, $\beta(G_l)$ is different from $\beta(G)$

$$\beta_{i,e_k}(G_l)=\begin{cases}\beta_{i,e_k}(G) & if\ l\notin P^i_{e_k}(d)\\ 0 & o.w.\end{cases}$$

- If the congested link doesn't belong to *i*-th path from source to destination, P^i, it will not affect packets going through those paths
  - It is zero otherwise

$$\beta_1(G)=\gamma_1\gamma_2 \qquad \beta_1(G_{l_1})=0$$
$$\beta_2(G)=\gamma_4\gamma_5 \qquad \beta_2(G_{l_1})=\beta_2(G)$$



Link Failure Model

- Training sequence is **A**
- $\mathbf{y}^l$: vector of symbols observed at the destination in *M* time slots with link l congested

$$y^l_{M\times 1}=A_{M\times N}\beta(G_l)_{N\times 1}$$

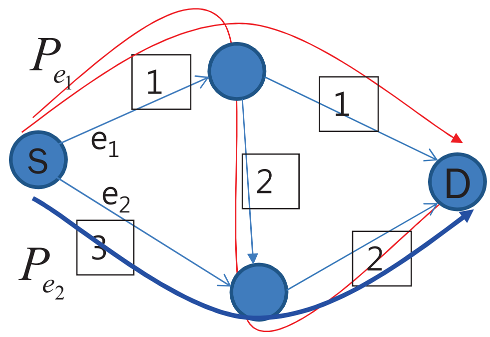
- Potential for identifying: received symbols change uniquely in response to link congestion

$$y_{M\times 1}\neq y^l_{M\times 1}$$
$$y^{l_1}_{M\times 1}\neq y^{l_2}_{M\times 1}$$

Example

$$\beta(G) = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} \qquad \begin{aligned} \beta_{1,e_1} &= 1 \times 1 = 1 \\ \beta_{2,e_1} &= 1 \times 2 \times 2 = 3 \\ \beta_{2,e_2} &= 3 \times 2 = 1 \end{aligned}$$

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 3 & 3 \end{bmatrix}$$



	--	e_1	e_2	ℓ_1	ℓ_2	ℓ_3
1st time slot	0	2	2	3	1	1
2nd time slot	2	3	1	0	1	3

Received symbols corresponding a single link failure

Theorem 2: Sufficient Conditions

- If Rank(**A**)= deg(S), and
  - for all P_{ek} set of paths between source and destination starting at e_k

$$\sum_{j=1}^{|P_{e_k}|} \xi_j \beta_{j,e_i} = 0 \Leftrightarrow \xi_j = 0 \, \forall j \qquad \text{(more next slide)}$$

then

$$A\beta(G) \neq A\beta(G_l) \quad \forall l \notin E$$

$$A\beta(G_{l_1}) \neq A\beta(G_{l_2}) \quad \forall l_1, l_2 \notin E$$

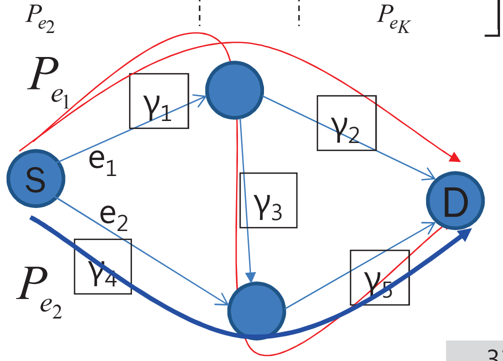
Theorem 2

- Condition $\sum_{j=1}^{|P_{e_k}|} \xi_j \beta_{j,e_i} = 0 \Leftrightarrow \xi_j = 0 \; \forall j$ means
  - For a set of paths having e_k in common, P_{ek} , NC coefficient of the paths are independent !

$$\left[\underbrace{\overset{\text{Independent}}{\beta_{1,e_1} \quad \beta_{2,e_1}} \quad \cdots \quad \beta_{N_1,e_1}}_{P_{e_1}} \;\vdots\; \underbrace{\overset{\text{Independent}}{\beta_{1,e_1}} \quad \cdots \quad \beta_{N_2,e_2}}_{P_{e_2}} \;\vdots\; \cdots \;\vdots\; \underbrace{\beta_{N_K,e_K}}_{P_{e_K}} \right]^t$$

$$\left.\begin{aligned} \beta_{1,e_1} &= \gamma_1\gamma_2 \\ \beta_{2,e_1} &= \gamma_1\gamma_3\gamma_5 \end{aligned}\right\} \text{independent}$$

$$\beta_{1,e_2} = \gamma_4\gamma_5$$



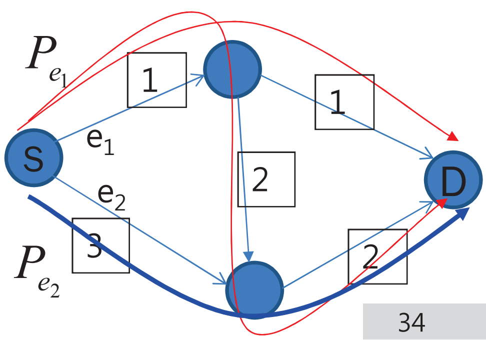
Example

❖ $$\beta(G) = \begin{bmatrix} 1 \\ 3 \\ \hdashline 1 \end{bmatrix}$$ (1 and 3: Independent)

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 3 & 3 \end{bmatrix} \quad Rank(A) = 2 = \deg(S)$$

$$\beta_{1,e_1} = 1 \times 1 = 1$$
$$\beta_{2,e_1} = 1 \times 2 \times 2 = 3$$
$$\beta_{2,e_2} = 3 \times 2 = 1$$

	--	e1	e2	ℓ_1	ℓ_2	ℓ_3
1st time slot	0	2	2	3	1	1
2nd time slot	2	3	1	0	1	3



Complexity/Speed

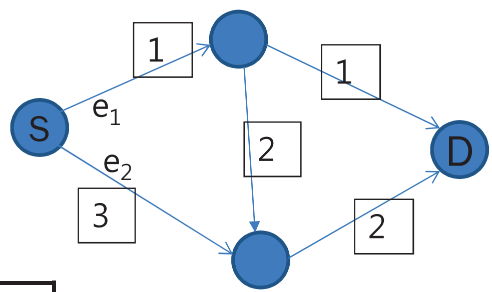
- First condition of Theorem 2:

$$\text{Rank}(A_{M\times N}) = \deg(S) \ \text{ implies } \ M \geq \deg(S)$$

  - In previous example M=2=deg(S)
- Number of time slots: at least the number of outgoing links of source
- Is it possible to decrease number of time slots? ➔ faster monitoring
- Possible by increasing number of bits in LNC coeff. ➔ more complexity

Example

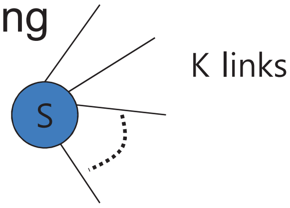
- q=3
- A=[1 1 4]



	--	e_1	e_2	ℓ_1	ℓ_2	ℓ_3
1st time slot	6	4	2	5	7	1

Theorem 3: Complexity/Speed tradeoff

- $N_i=|P^i|$
- *q* bits per symbol are used in network coding
- *M* number of (desired) time slots
- Let Z={1,2,…,K}
- *K* degree of source
- ***Z***$_M$: collection of all partitions of *Z* with size *M*



$$Z_M = \{\{H_1, H_2, ..., H_M\} \mid \bigcup_{i=1}^{M} H_i = Z, H_i \cap H_j = \Phi\}$$

- *K*=3, *M*=2 ➔ Z={1,2,3}
- Z_M={ {{1,2},{3}} , {{1,3},{2}} , {{2,3},{3}} }

Journal Club Meeting, June 27, 2013

Theorem 3: Complexity/speed tradeoff

- Network is 1-identifiable if

$$q \geq \min_{\{H_i, i=1,...,M\} \in Z_M} \max_i \sum_{j \in H_i} N_j$$

Rank(**A**)=M

Theorem 3 provides a tradeoff between number of time slots for training sequence (speed of the method) and size of network coding coefficient (complexity) to make a network *G(V; E)* identifiable.