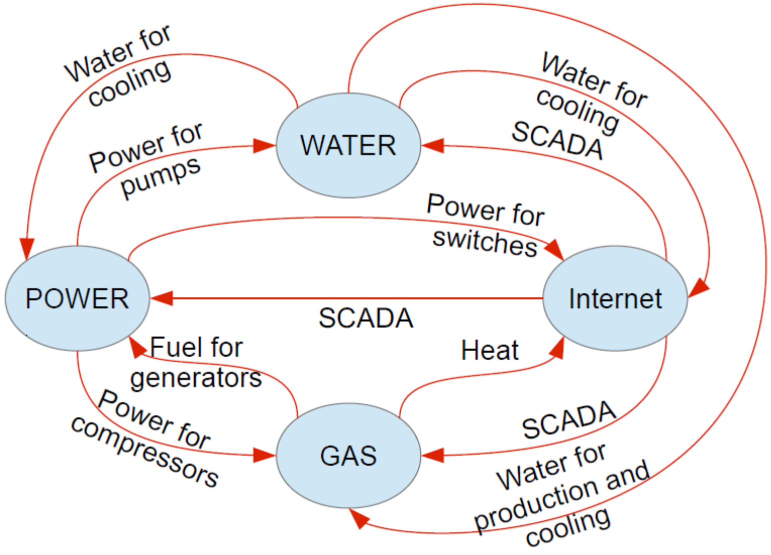
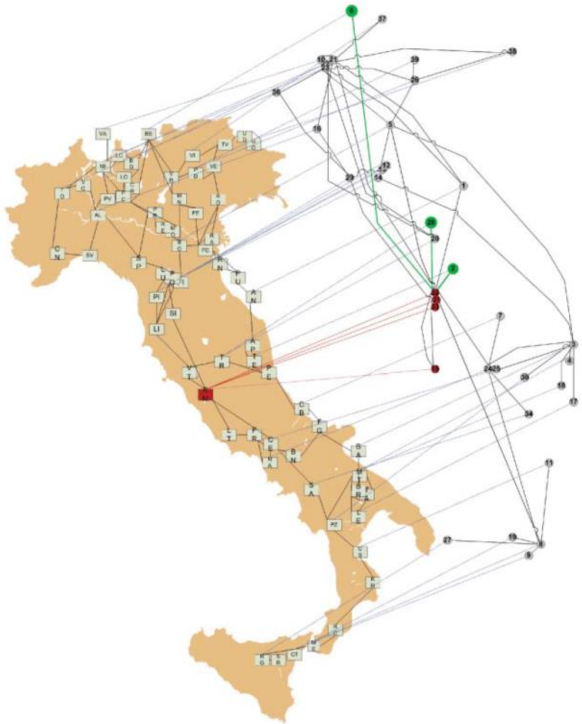
Introduction to joint projects for the courses of Intensive Computation and Performance of Computer Networks

Novella Bartolini and Annalisa Massini

Damage assessment and recovery after network failures

Massive network failures in networks may derive from single failures





© S. Buldyrev et al., Nature, Letters, Vol 464, 2010

Failure of nodes in one network causes failure of nodes in a second network

Supervisory Control And Data Acquisition (SCADA systems) cause interdependency

communication network – other infrastructures

Structural heterogeneity

Different behaviors of propagation

Problem Setting

Supply Graph *G*

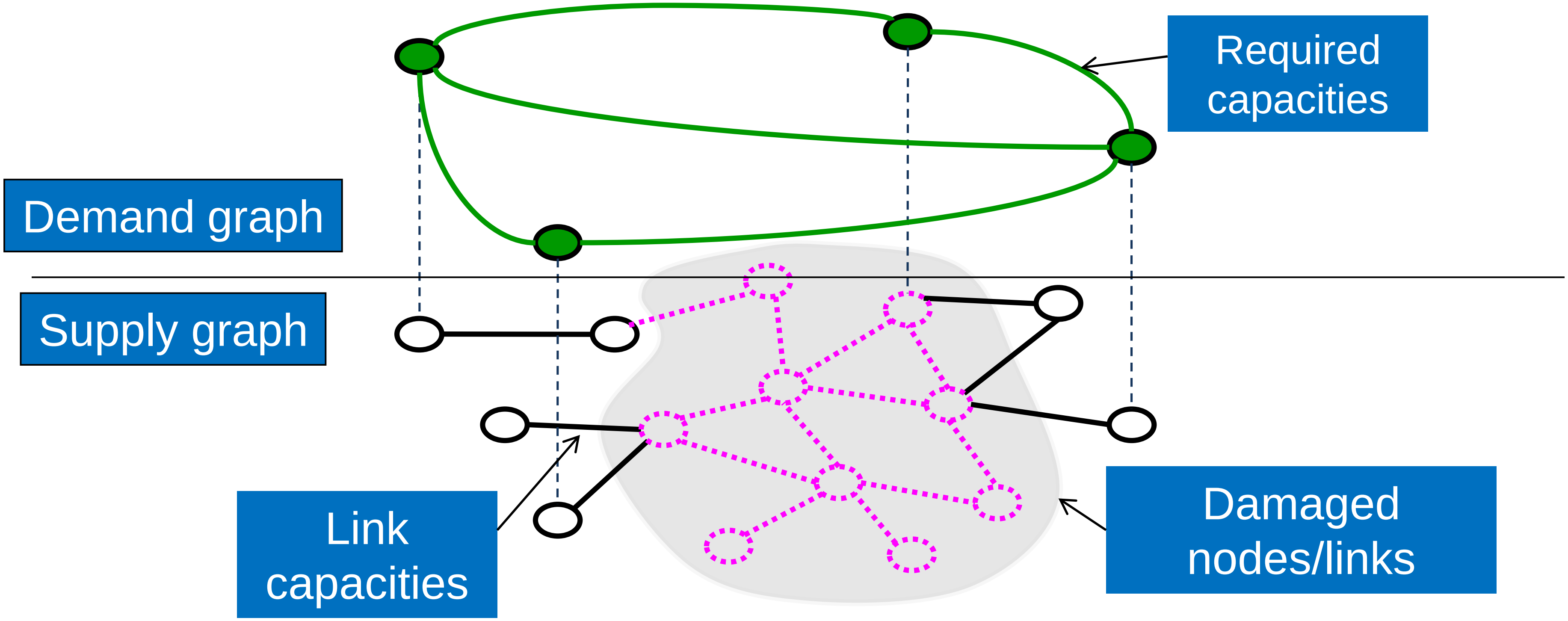
- Damaged communications network

Demand Graph *H*

- Flows with required capacity for mission critical applications

Goal

- Make lowest cost repairs (restorations) in *G* to serve all flows in *H*



Required capacities

Demand graph

Supply graph

Link capacities

Damaged nodes/links

Network failures

Network management under failures

- Analysis and design (models of failure propagation, network engineering)
- Assessment (monitoring and network tomography)
- Recovery (algorithms for service restoration)

Related funded projects and collaborations:

ARL (Army Research Lab)

DTRA (Defense and Threat Reduction Agency)

Collaborations with Penn State University and IBM

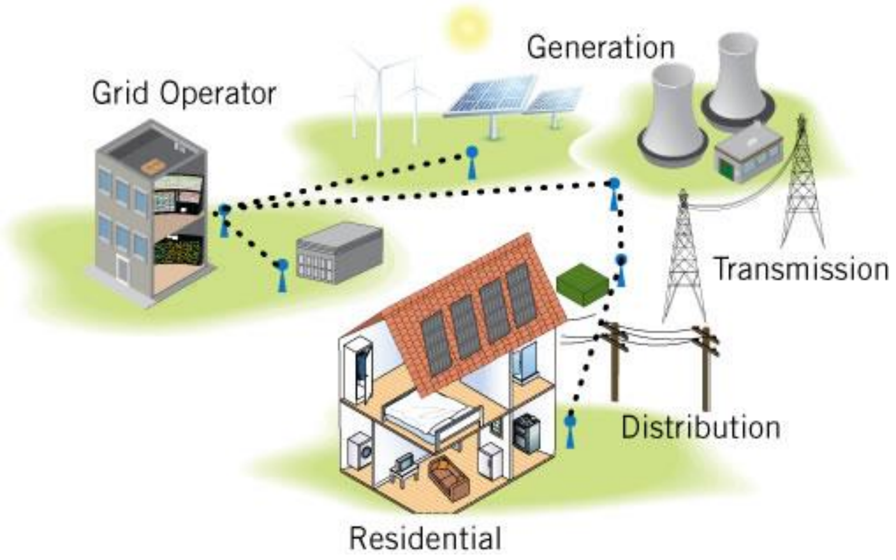
Cascading failures

[1] H. Khamfroush, N. Bartolini, T. La Porta, A. Swami, J. Dillman,
On Propagation of Phenomena in Interdependent Networks,
in *IEEE Transactions on Network Science and Engineering*, Vol. 3, n. 4, July 2016.

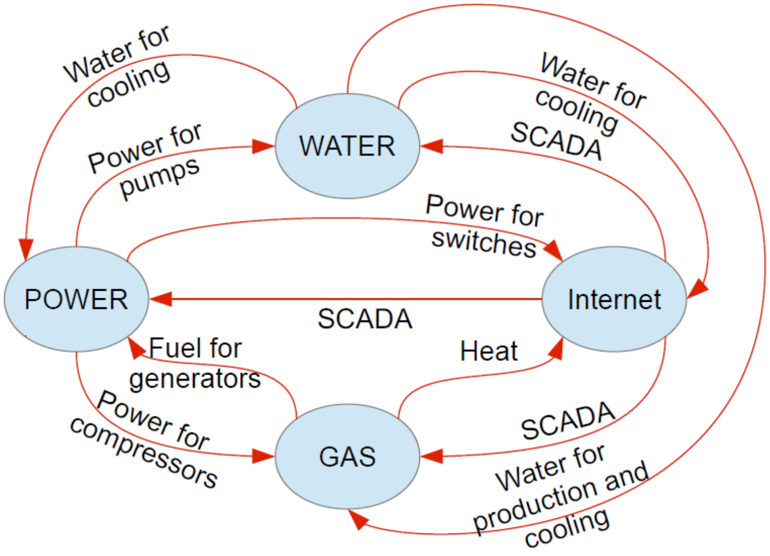
Large scale interdependent networks

- Interdependent networks: functionality or performance of one network depends on the other

Internet controls power grid &
grid provides power for the Internet



Massive network failures in networks



WATER

POWER

Internet

GAS

Water for cooling

Power for pumps

Water for cooling

SCADA

Power for switches

SCADA

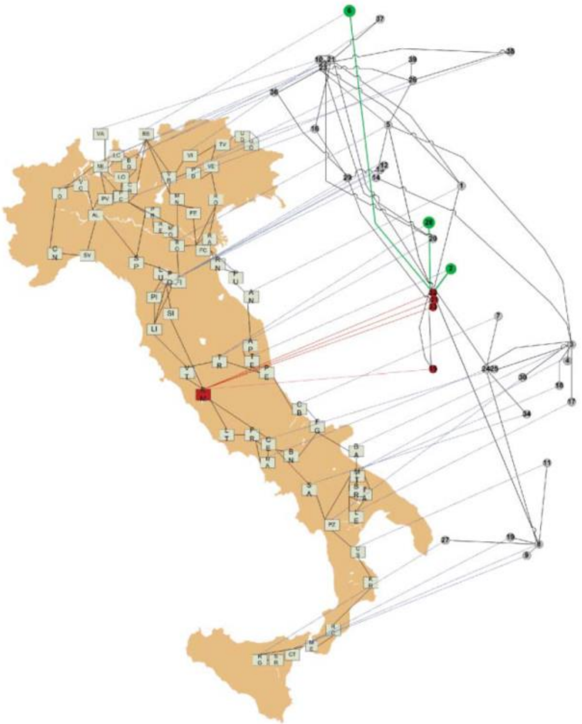
Fuel for generators

Heat

Power for compressors

SCADA

Water for production and cooling



© S. Buldyrev et al., Nature, Letters, Vol 464, 2010

Failure of nodes in one network causes failure of nodes in a second network

Supervisory Control And Data Acquisition (SCADA systems) cause interdependency

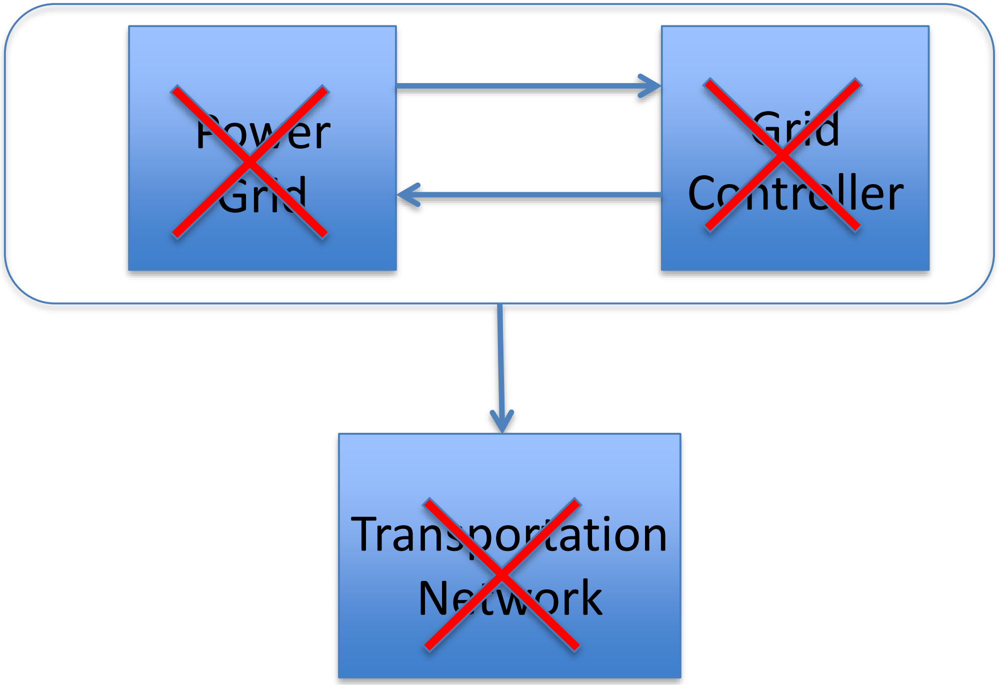
communication network – other infrastructures

Structural heterogeneity

Different behaviors of propagation

Motivation

- Blackout in Italy, Sep 2003 : Power outage affected all Italy
- 56 million people have been affected

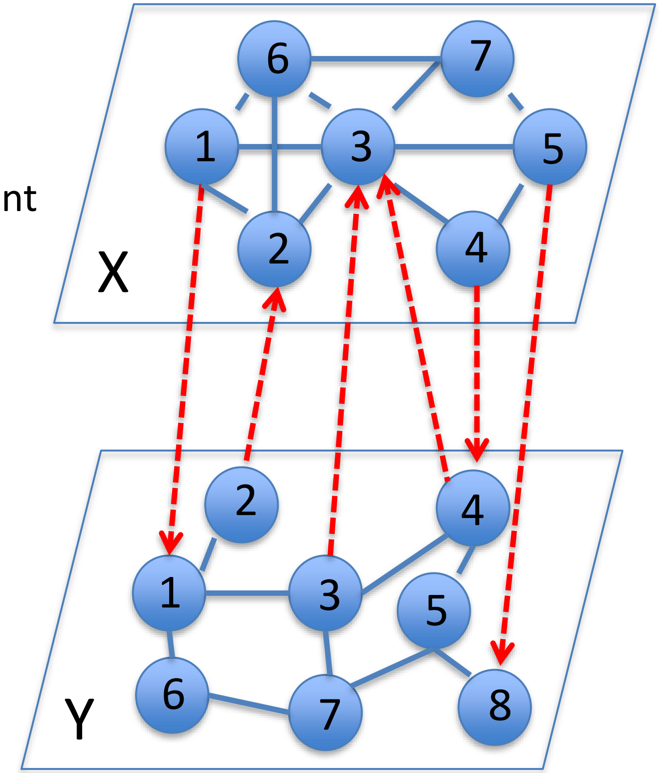


Cascading failures

- Two inter-dependent networks X and Y with respectively, n_x and n_y nodes
  - Red links represent inter-connectivity and blue links represent intra-connectivity links
- Given the initial spreaders set
  - Calculate the probability of transition into a new state
  - Expected time to full spread or end of the propagation

Example: For node *3* of network *X*

- Set of intra-connection ={1,2,3,5,6,7} of *X*
- Set of inter-connection={3,4} of *Y*



Problem 1: characterize the propagation, control the speed of propagation

Problem 2: design robust networks (with failure detection capability and slow propagation)

Network tomography

[1] Ting He, Novella Bartolini, Hana Khamfroush, InJung Kim, Liang Ma, Tom La Porta,
Service Placement for Detecting and Localizing Failures Using End-to-End Observations,
in *Proceedings of the 36th IEEE International Conference on Distributed Computing Systems (IEEE ICDCS 2016)*

[2] N. Bartolini, T. He, H. Khamfroush,
Fundamental Limits of Failure Identifiability by Boolean Network Tomography,
in *IEEE Proceedings of the International Conference on Computer Communications (IEEE INFOCOM 2017)*

Network tomography

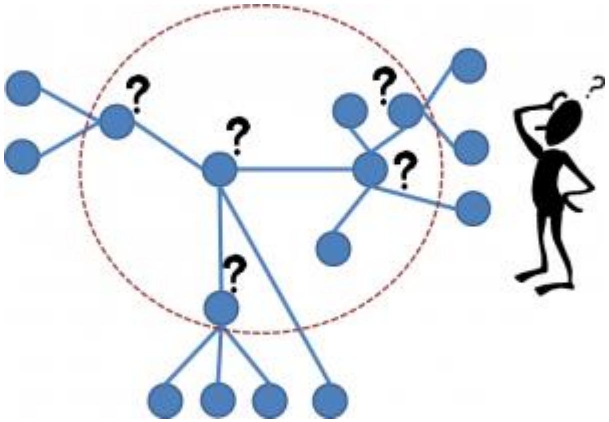
Network Tomography:

Inferring internal network state through external, end-to-end measurements

Relevance

Knowledge of the network state is important

- Prompt intervention after failure
- Efficient Routing
- Resource Allocation
- Balancing network loads
- QoS measurement: service degradation



Challenges

Large and costly overheads due to active probing

Problem 1: Optimal monitor placement for detecting and localizing failures

Problem 2: Minimize number of monitoring paths

Problem 3: Maximize identifiability of failures

Problem 4: Design new network topologies with maximum identifiability of failures

Sensor and actuator networks

(drones + terrestrial robots + sensor networks)

Related funded projects:

NATO Science for Peace and Security G4936,
Hybrid Sensor Networks for Emergency Critical Scenarios
(2015-2018, in collaboration with GJU and MS&T)

PSU seed project,
Digital innovation in food security using a 28,000 farmer living lab in Kenya

Monitoring drones

Sensors can be mounted on drones.
In this case they are typically complex sensing devices interfaced with artificial intelligence for image processing, event recognition.

Why a network and not a single drone doing all the work?

Amatrice – Italy (2016)

Why a network and not a single drone doing all the work?

In the aftermath of a catastrophe, drones are used to find people, provide medicines to inaccessible and possibly unknown locations.

The intervention must be fast, as it may save lives.

The battery of the drone, especially with payload, ensures a limited flight time.

Better to use multiple coordinated drones, which autonomously spread through the area.

Why a network and not a single drone doing all the work?

- The use of a squad in inaccessible terrains is also motivated by the limited supplies available on site

Examples:

low/high temperatures (imagine you are monitoring a glacier),

absence of roads,

absence of connectivity…

Current work on Sensor and Actuator Networks

Field crops at Penn State

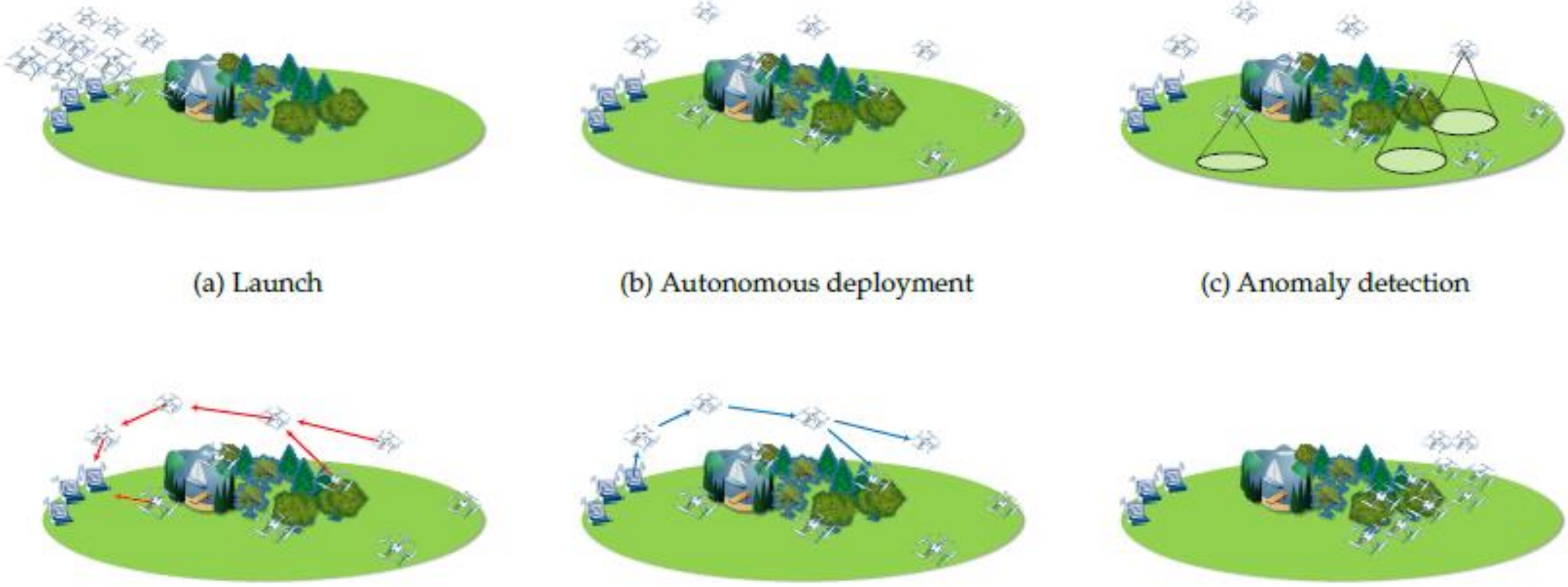
Current work on Sensor and Actuator Networks

Farms in the Philippines

Current work on Sensor and Actuator Networks

Farms in Uganda

Current work on Sensor and Actuator Networks



(a) Launch

(b) Autonomous deployment

(c) Anomaly detection

(d) Transmission to sink

(e) Task assignment

(f) Task execution

Research challenges

- Different concept of coverage to be optimized! Dynamic coverage: a point is covered if it is traversed, or if it is explored. There may be deadlines.

- Flight at different heigths cause different sensing capabilities. The propeller wings cause noise in the measurements. Height

- Battery limitations are rigid, you can recharge the device but you cannot let it drop!

-> Analytical formulation of optimization problems, algorithmic solutions

Boolean Network Tomography

N. Bartolini, A. Massini, et al.

Fundamental Limits of Failure Identifiability by Boolean Network Tomography,

Outline

- Motivation
- Network Tomography
- Definitions
- Problem Formulation
- General Network Monitoring Bounds
- Service Network Monitoring Bounds
- Performance Evaluation

Motivation

- Identifying the state of network nodes is beneficial for many functions in network management
  - Performance analysis
  - Route selection
  - Network recovery

- Direct measurement is not always available due to large traffic overhead, access control, etc.

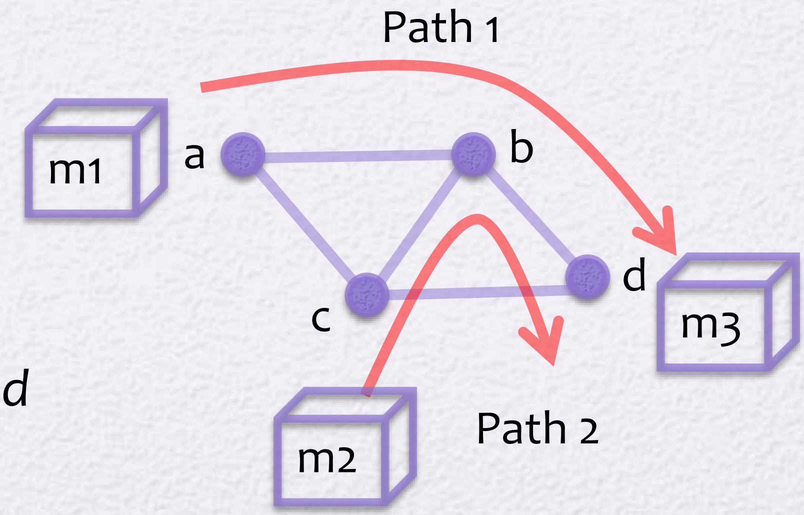
- Built-in monitoring may fail detecting failures caused by misconfigured/unanticipated interactions between network layers (silent failures)

One solution: Network Tomography

Boolean Network Tomography (BNT)

- Diagnose the health of network elements from the health of end-to-end communications perceived between measurement points

- Node states can be measured indirectly via monitoring paths



Path 1 & path 2 fail: can't localize
Path1 fails, path 2 working: link ab failed

Our Problem Setup

- Network is modeled as undirected graph $G=(V,E)$, V representing nodes and E representing links

- Failure set, i.e. set of failed nodes: $F \subseteq V$

- Total number of nodes: n

- Nodes states can be measured indirectly only by monitoring paths

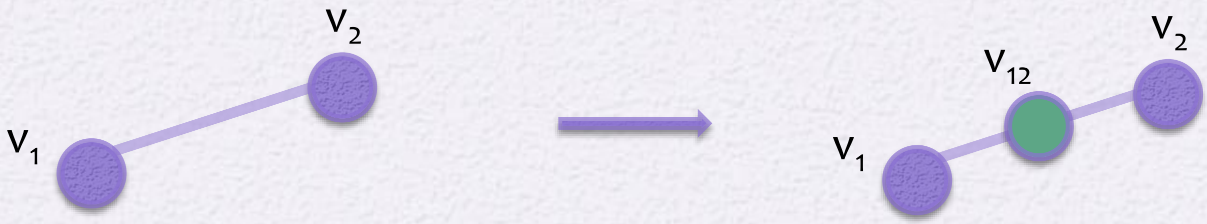
- Set of monitoring paths: $P=\{p1,p2,\ldots,pm\}$

- The state of a path is normal if all traversed nodes are in normal state

Our Problem Setup

- Failure set, i.e. set of failed nodes: $F \subseteq V$

  Notice that we focus on the failure of nodes only, as links can be modeled as virtual nodes.



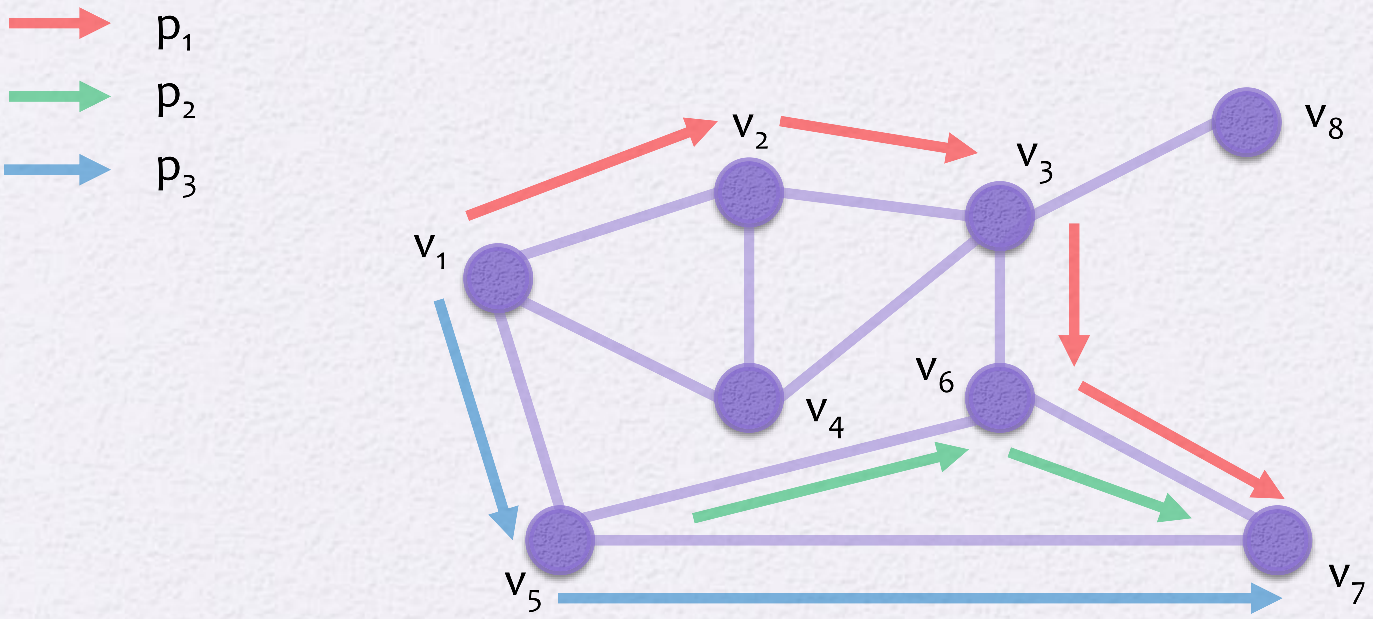
Node v_{12} represents the status of link (v_1, v_2)

Our Problem Setup

- Incident set of v_i : set of paths affected by the failure of node v_i noted by P_{v_i}
- Incident set of paths of a failure set *F*: $P_F \triangleq \cup_{v_i \in F} P_{v_i}$
- Test matrix T is an $m \times n$ matrix, where $T|_{i,j} = 1$

  if $v_j \in p_i$ and zero otherwise
- The *j-th* column of T denoted with $b(v_j) \triangleq T|_{*,j}$ is the characteristic vector of P_{v_j} and called binary encoding of v_j

Test matrix T



$$T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Identifiability Definition

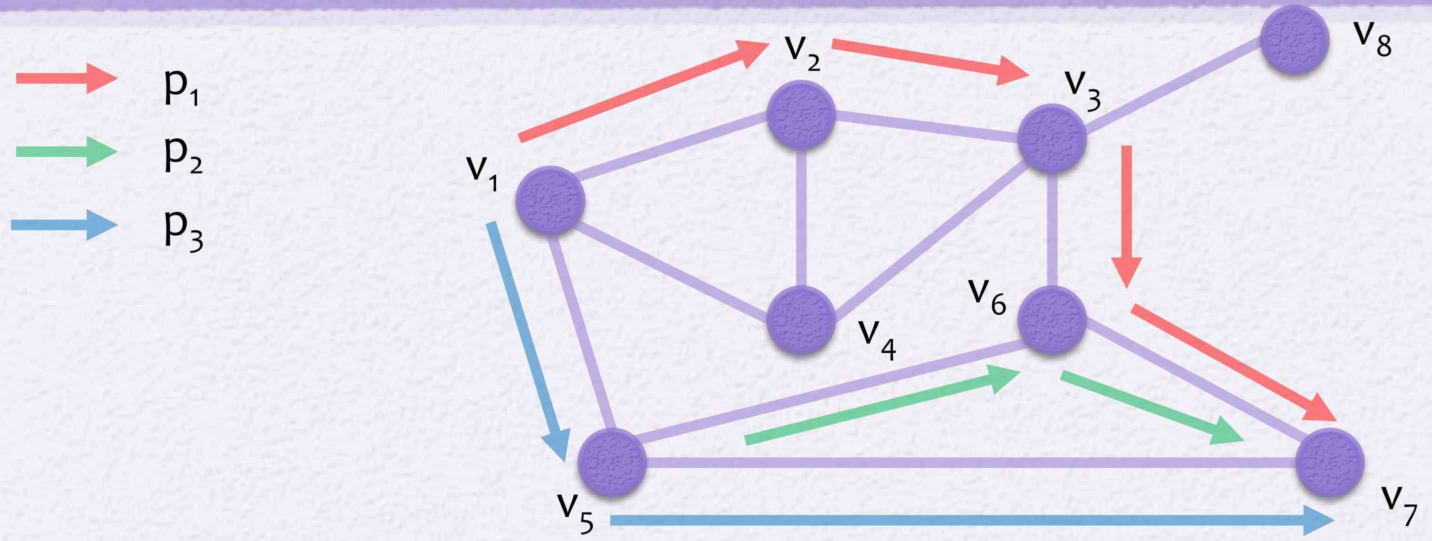
Definition *Given a set of monitoring paths P and a node $v_j \in V$, v_j is k*-identifiable *with respect to (wrt) P if for any failure sets F_1 and F_2 such that $F_1 \cap \{v_j\} \neq F_2 \cap \{v_j\}$, and $|F_i| \leq k$ ($i \in \{1, 2\}$),*

$$\bigvee_{v_i \in F_1} b(v_i) \neq \bigvee_{v_z \in F_2} b(v_z)$$

where with "$\bigvee$" we refer to the element-wise logical OR.

Definition *A node v_i is 1-identifiable wrt P if and only if $b(v_i) \neq \mathbf{0}$, and $\forall v_j \neq v_i$, $b(v_j) \neq b(v_i)$, i.e., its binary encoding is not null and not identical with that of any other node.*

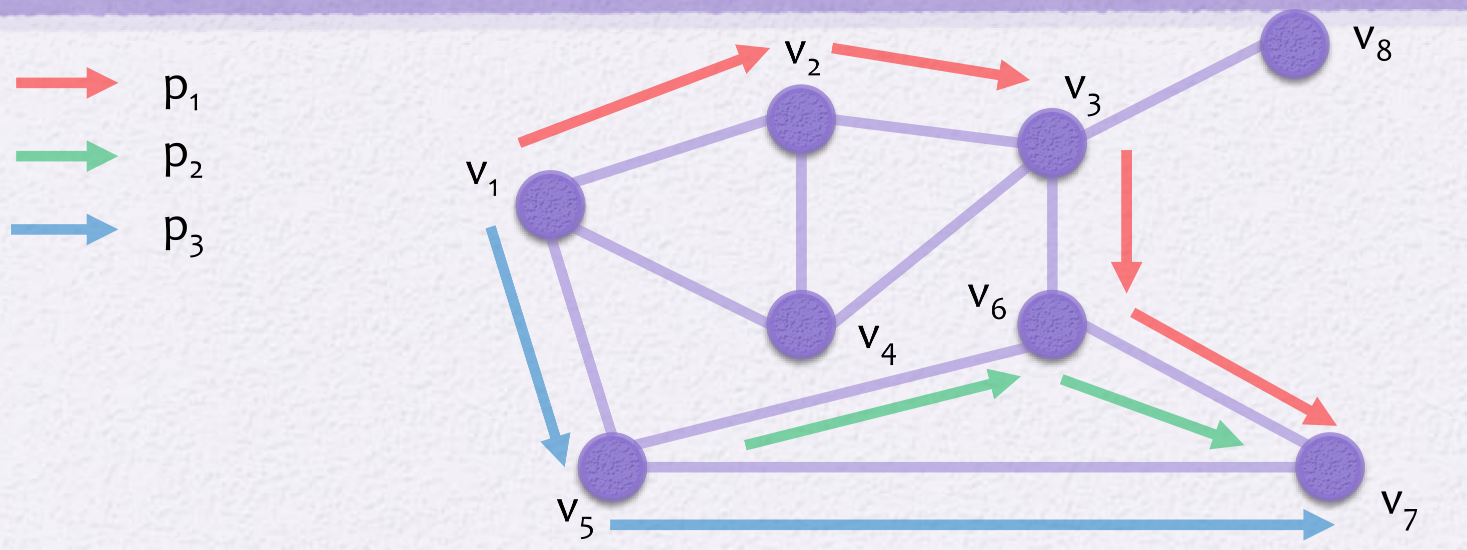
Test matrix T



$$T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Which nodes are 1-identifiable?

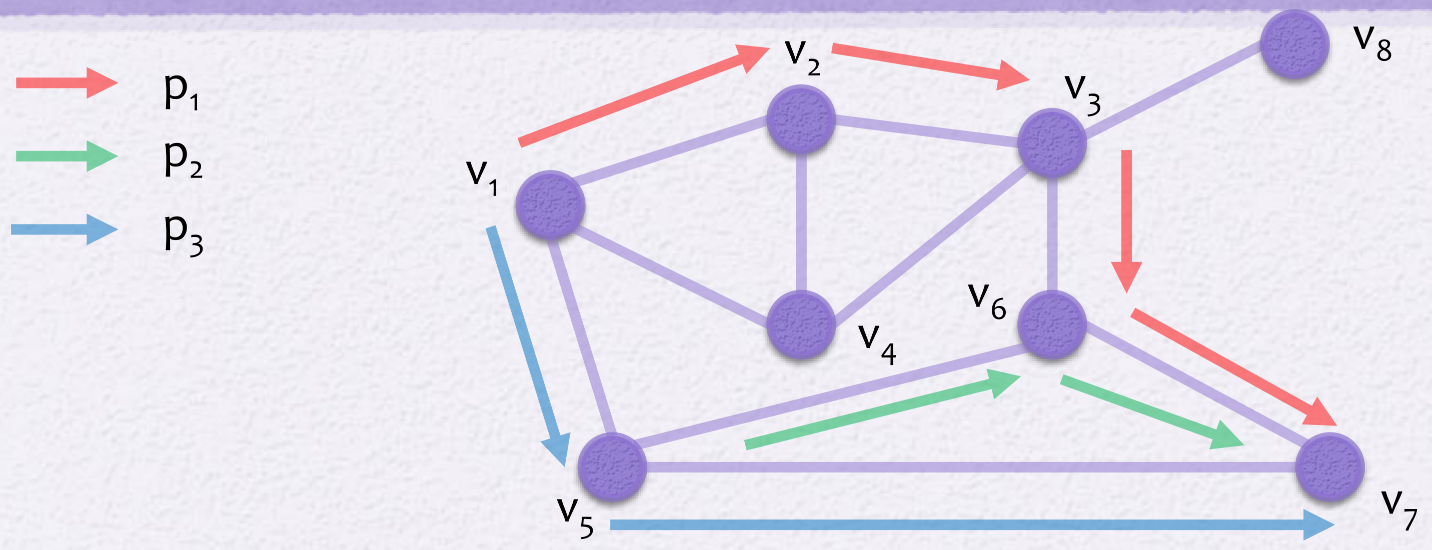
Test matrix T



$$T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Which nodes are 1-identifiable? v_1, v_5, v_6, v_7

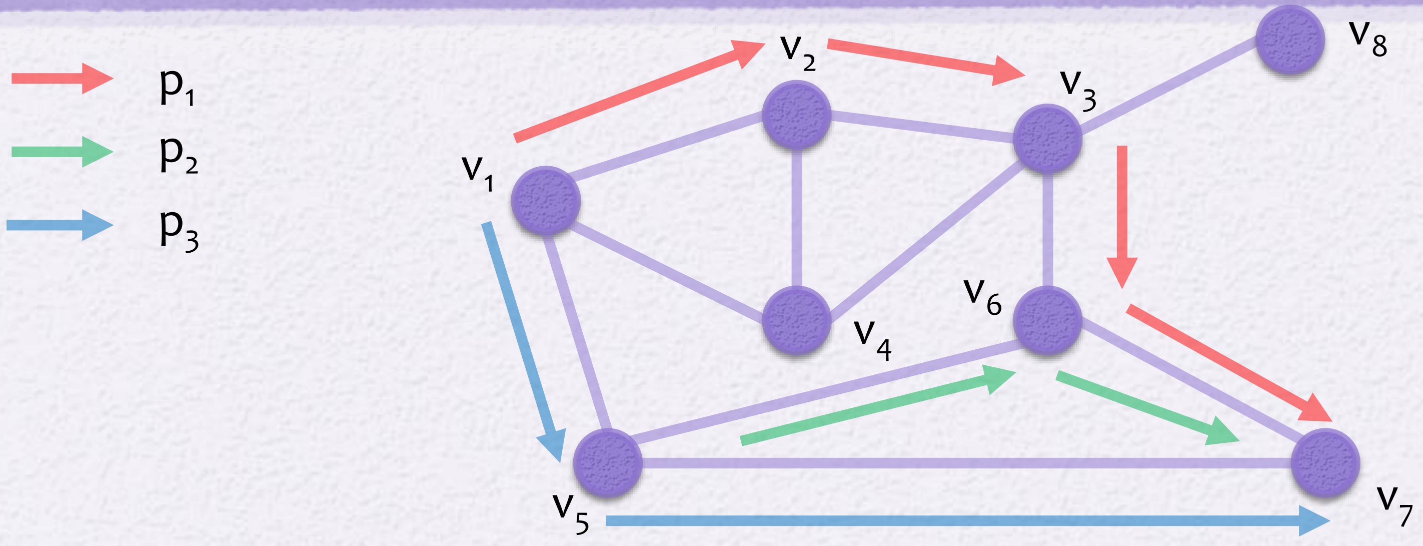
Test matrix T



$$T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Why is v_4 not identifiable?

Test matrix T

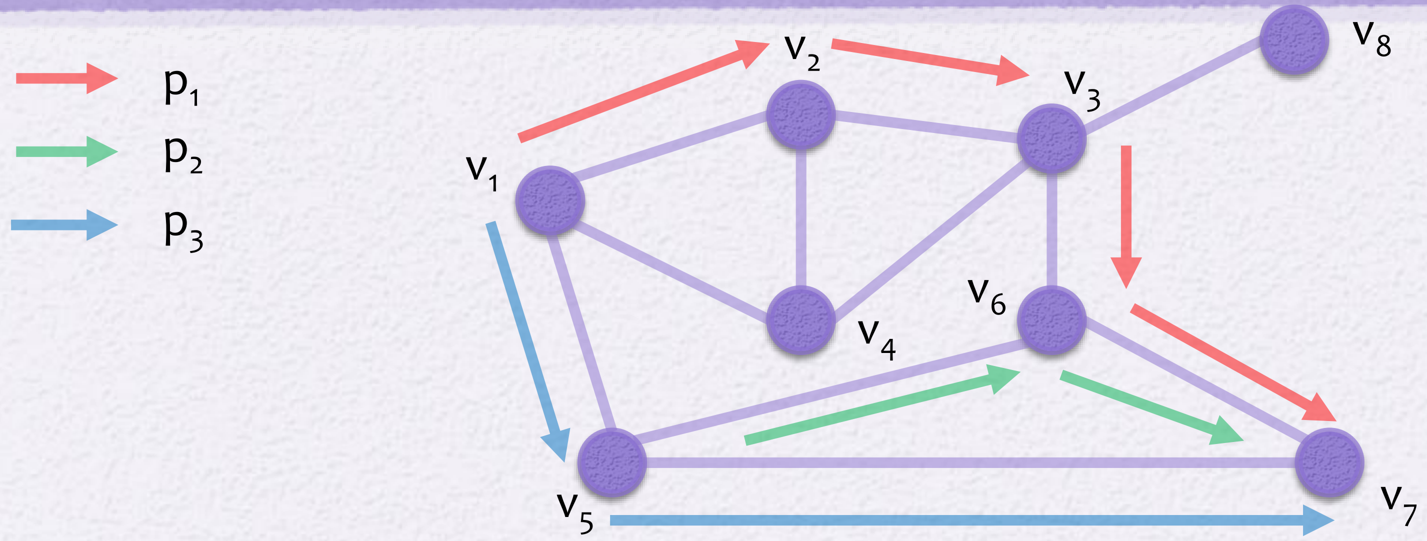


$$T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Why is v_4 not identifiable?

It is not even traversed by any path! Same as v_8.

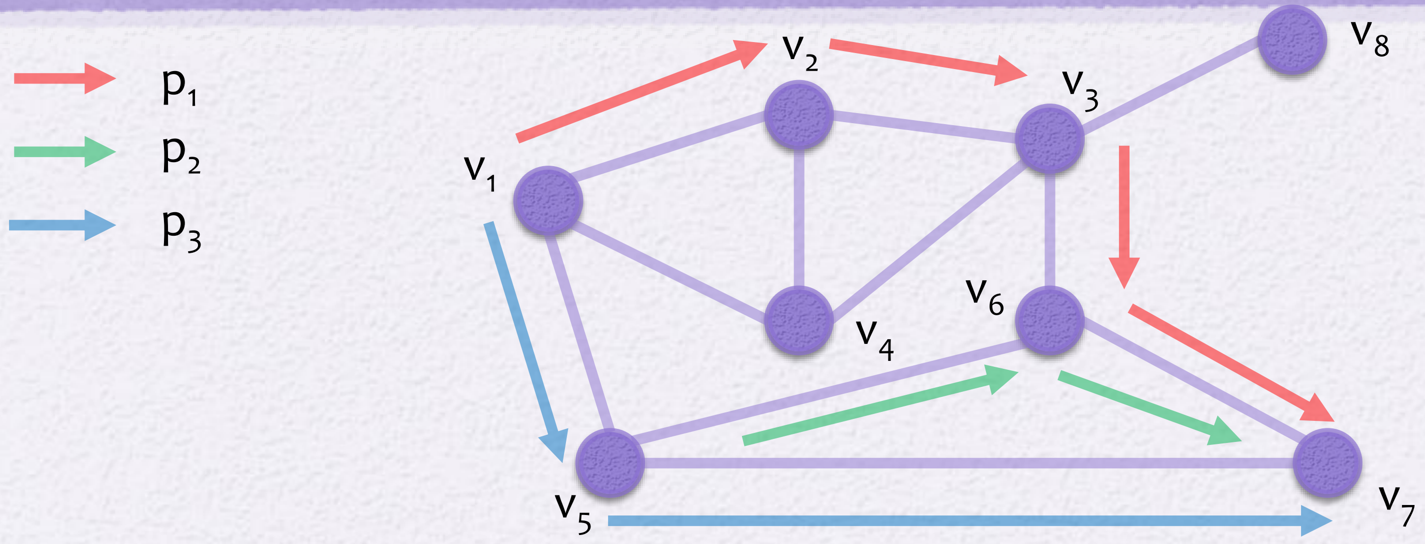
Test matrix T



$$T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Why are v_2 and v_3 not identifiable?

Test matrix T

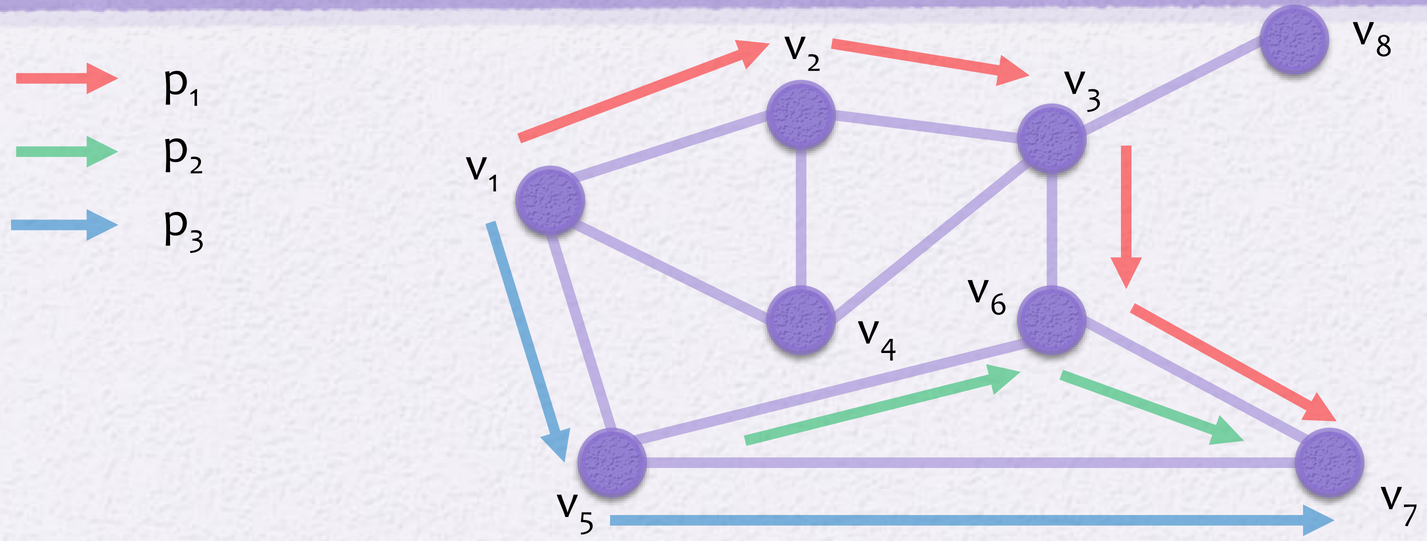


$$T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Why are v_2 and v_3 not identifiable?

They have the same Boolean encoding! Whatever failure occurs, we cannot distinguish v_2 from v_3.

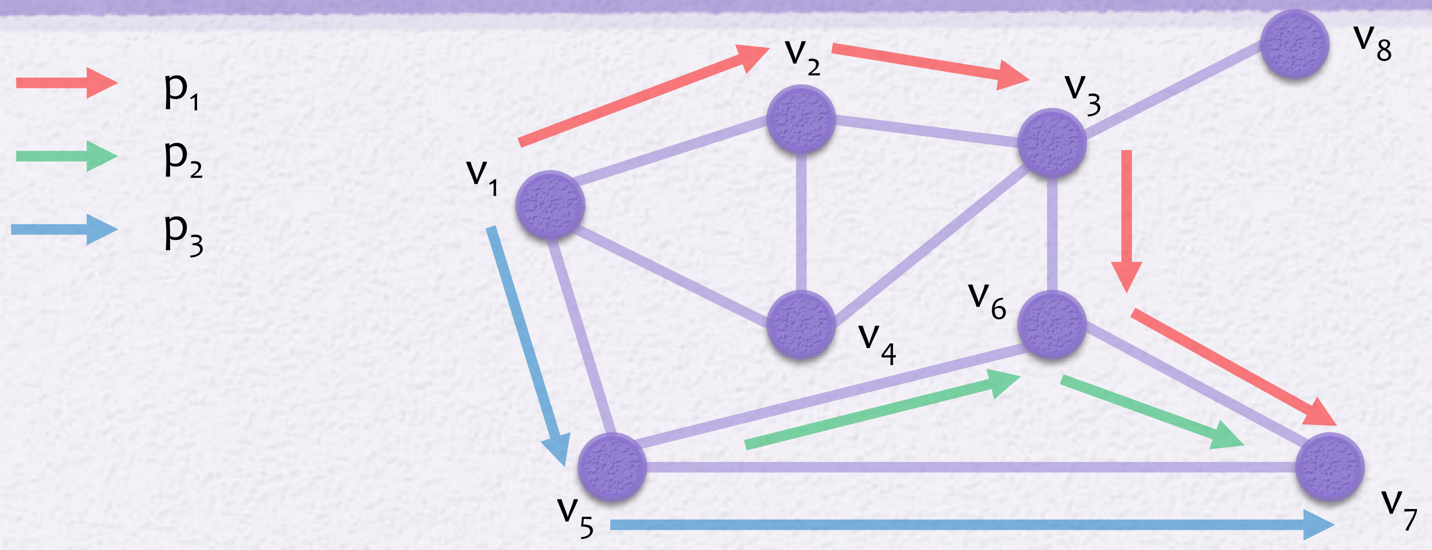
Test matrix T



$$T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Which nodes are 2-identifiable?

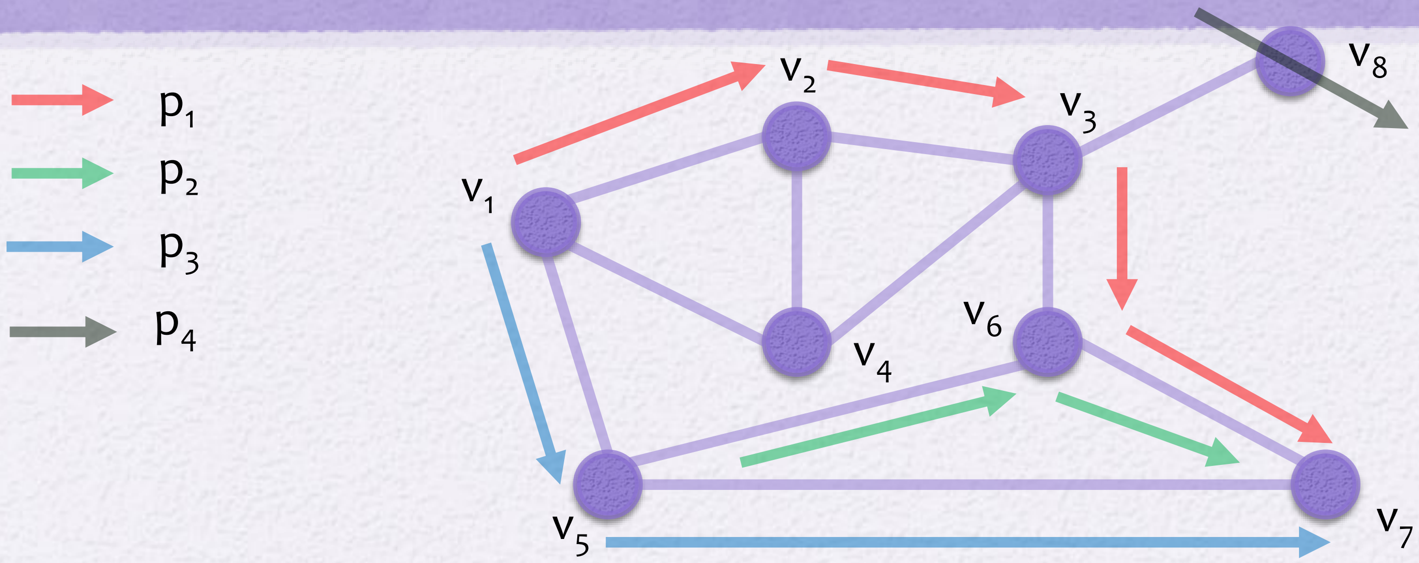
Test matrix T



$$T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Which nodes are 2-identifiable? None of them!

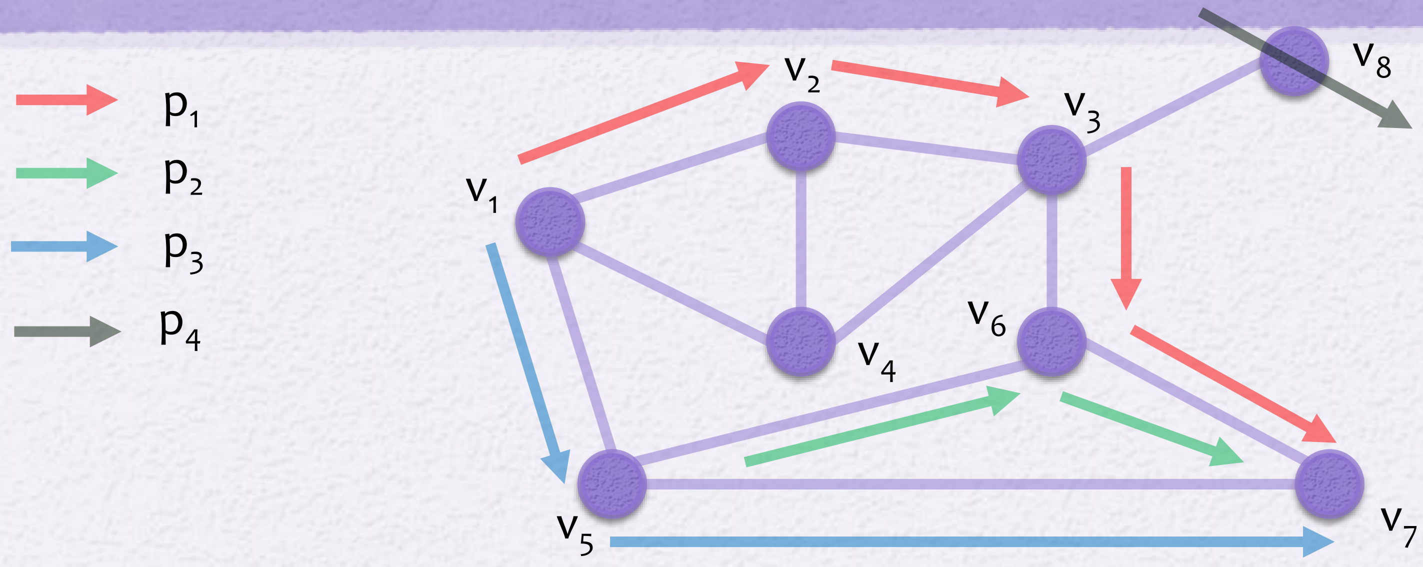
Test matrix T



$$T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Which nodes are 2-identifiable?

Test matrix T



$$T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Which nodes are 2-identifiable? v_8

Problem Definition

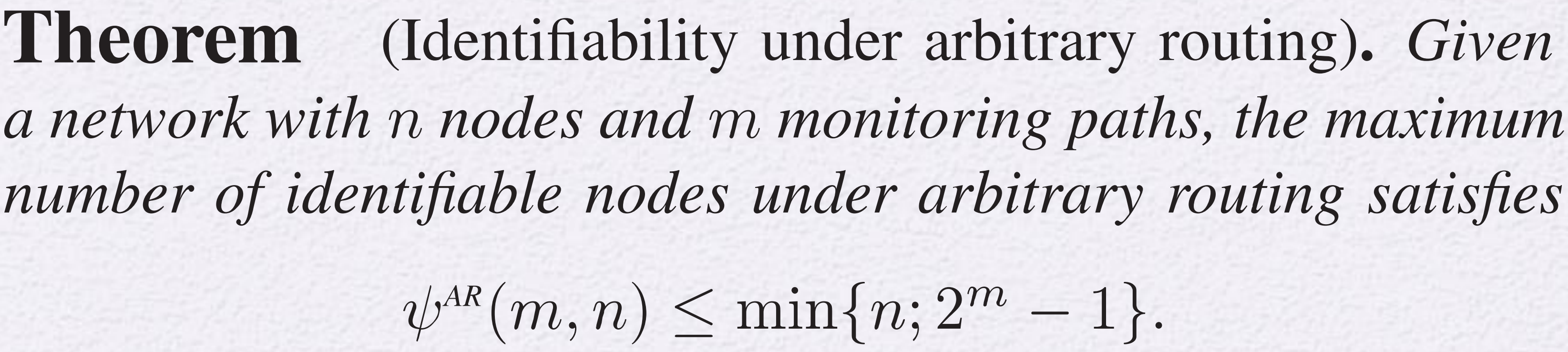
- Given a collection of candidate path sets *P* under all possible designs, how well can we monitor the network using path measurement and which design is the best?

- Monitoring performance is measured by the number of nodes that are *k-identifiable* w.r.t *P*

- The optimal solution is hard due to the exponential number of path sets

- We focus on bounding the number of *1-identifiable nodes, since the upper bound on 1-identifiable would be an upper bound on k-identifiable as well*

General Network Monitoring Arbitrary routing



Theorem (Identifiability under arbitrary routing). *Given a network with n nodes and m monitoring paths, the maximum number of identifiable nodes under arbitrary routing satisfies*

$$\psi^{AR}(m,n) \leq \min\{n; 2^m - 1\}.$$

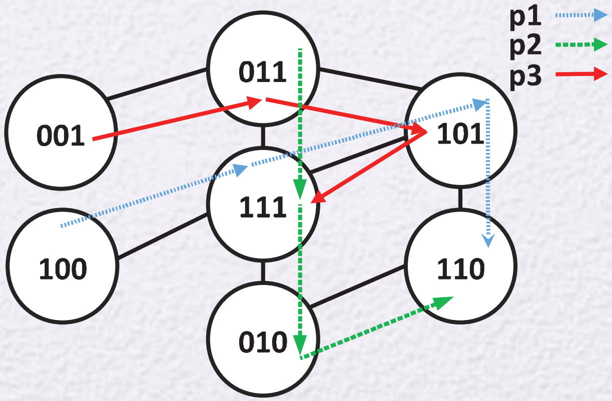
General Network Monitoring Arbitrary routing

Theorem (Identifiability under arbitrary routing). *Given a network with n nodes and m monitoring paths, the maximum number of identifiable nodes under arbitrary routing satisfies*

$$\psi^{AR}(m, n) \leq \min\{n; 2^m - 1\}.$$

The bound is tight since we can construct a topology with m monitoring paths that meets this bound:

1. *Take up to 2^m nodes*
2. *Give binary enumeration*
3. *Construct paths*
4. *Create the edges of the graph*



General Network Monitoring Consistent routing

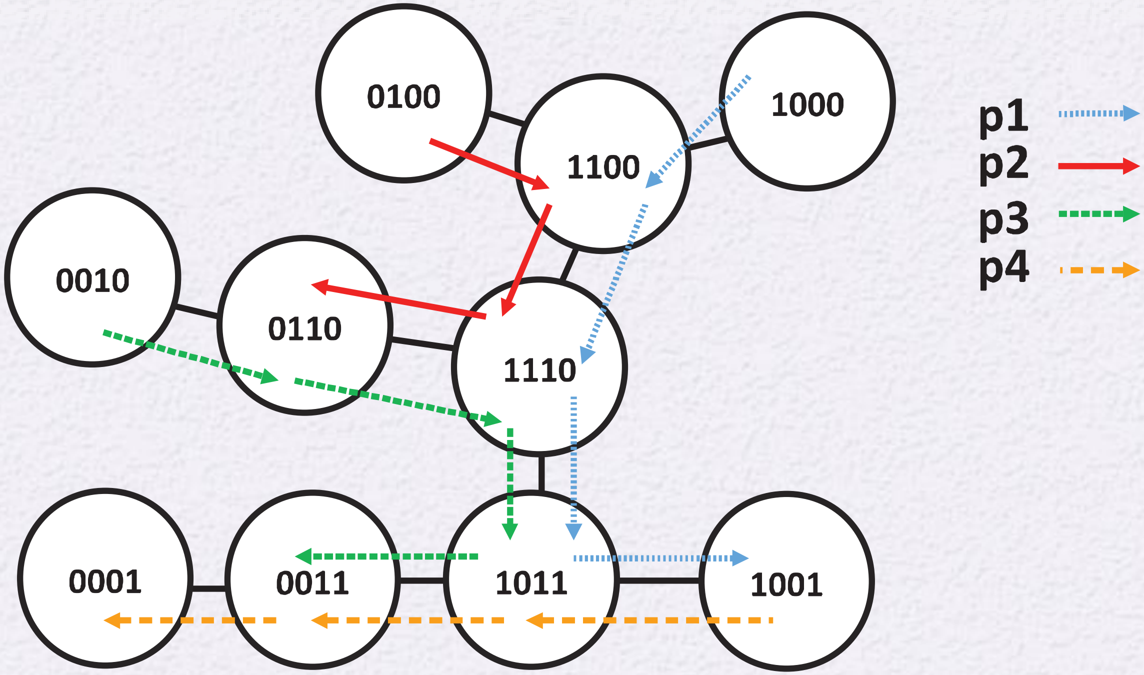
II. Consistent routing

Definition *A set of paths P is* consistent *if $\forall p,\ p' \in P$ and any two nodes u and v traversed by both paths (if any), p and p' follow the same sub-path between u and v.*

Definition

We define the *path matrix* of $\hat{p}_i$ as a binary matrix $M(\hat{p}_i)$, in which each row is the binary encoding of a node on the path, and rows are sorted according to the sequence $\hat{p}_i$. Notice that by definition $M(\hat{p}_i)|_{*,i}$ has only ones, i.e., $M(\hat{p}_i)|_{r,i} = 1,\ \forall r$.

Example of path matrix



$$M(\hat{p}_3) = \begin{array}{c|cccc} \text{flips} & b_1 & b_2 & b_3 & b_4 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 2 & 1 & 1 & 1 & 0 \\ 3 & 1 & 0 & 1 & 1 \\ 4 & 0 & 0 & 1 & 1 \end{array}$$

General Network Monitoring
Consistent routing

Lemma *Under the assumption of consistent routing, all the columns in all the path matrices have consecutive ones.*

Lemma *Given* $m = |P| > 1$ *consistent routing paths, whose length is at most* d^* *(in number of nodes), the maximum number of different encodings in the rows of* $M(\hat{p}_i)$ *is equal to* $\min\{2 \cdot (m-1), d^*\}, \forall p_i \in P$.

General Network Monitoring Consistent routing

Theorem (Identifiability with consistent routing). *Given n nodes, and $m > 1$ consistent routing paths of length at most d^* (in number of nodes), the maximum number of identifiable nodes satisfies:*

$$\psi^{CR}(m,n,d^*) \leq \min\left\{\sum_{i=1}^{i_{max}}\binom{m}{i} + \left\lfloor \frac{N_{max} - \sum_{i=1}^{i_{max}} i \cdot \binom{m}{i}}{i_{max}+1} \right\rfloor ; n\right\},$$

where $i_{max} = \max\{k \mid \sum_{i=1}^{k} i \cdot \binom{m}{i} \leq N_{max}\}$,
and $N_{max} = m \cdot \min\{2 \cdot (m-1); d^*\}$.

General Network Monitoring Consistent routing

Proof

- Each *identifiable* node must have a unique encoding
- For every path matrix, we have 2*(m-1) possible different encodings, so totally m*min{2*(m-1), d*}
- We are counting multiple times the nodes that appear in multiple path matrices
- If encoding b has k digits equals to 1, then b appears among the rows of k different path matrices

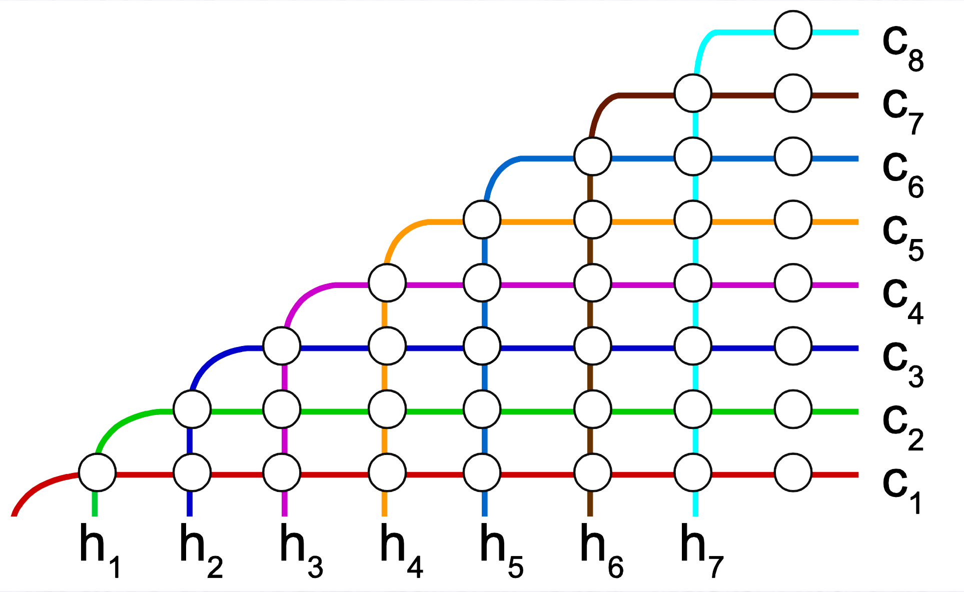
General Network Monitoring Consistent routing

Proof

- Number of distinct encoding is maximized when the number of duplicate encodings is minimized, therefore their number of ones is minimized
- Minimum number of duplicate is achieved when we have $\binom{m}{1}$ different encodings with only one digit equal to one, $2\binom{m}{2}$ with two digits equal to one appearing in two path matrices and so forth until total number of encodings is equal to N_{max}

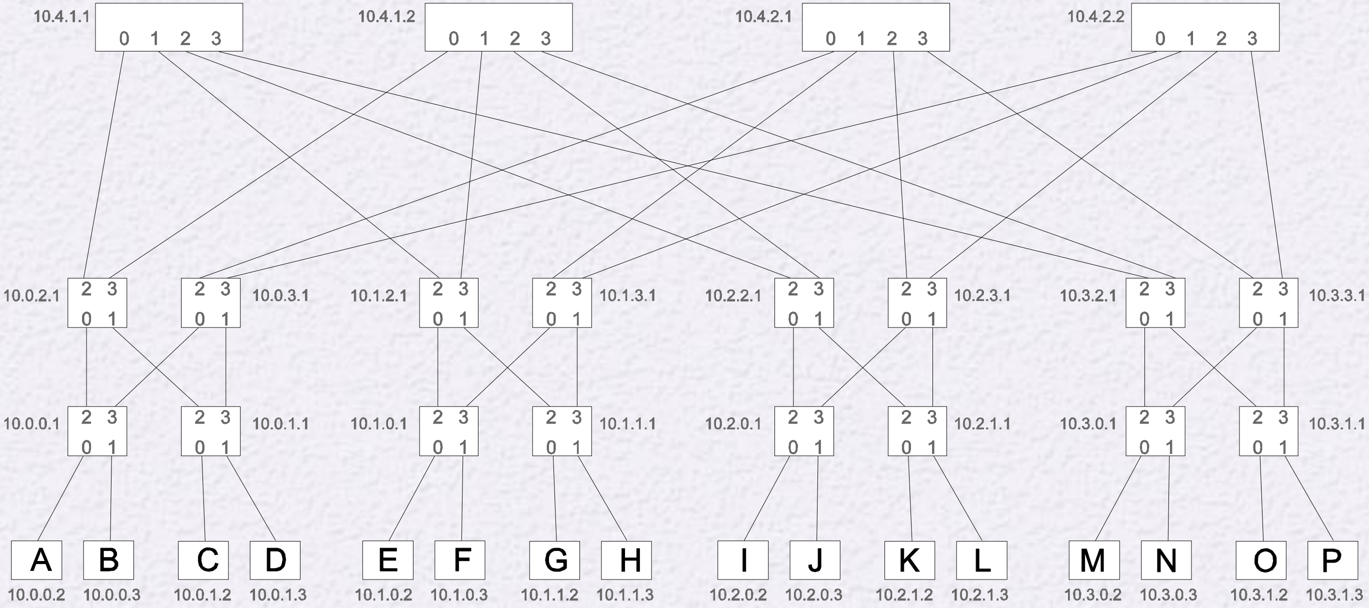
General Network Monitoring
Consistent routing

Tightness of the bound on number of identifiable nodes under consistent routing



With n=36 nodes, m=8 monitoring paths of maximum length d^*=8, we have N_{max}=min{112,64}=64, i_{max}=2, and
ψ^{cr}= $\binom{8}{1} + \binom{8}{2} + \left\lfloor \frac{0}{3} \right\rfloor$=36.

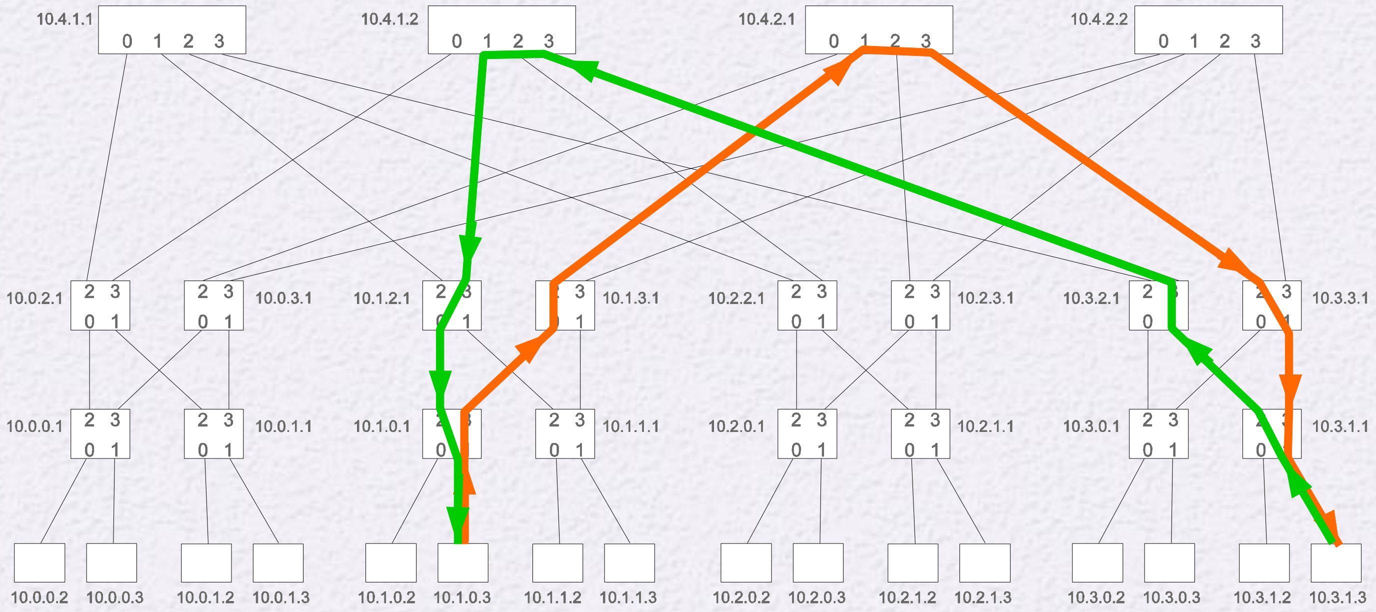
General Network Monitoring
The Case of Half-Consistent Routing



Fat-tree topology (common in data centers), where we assume the routing scheme based on IP addressing of clients and switches as described in

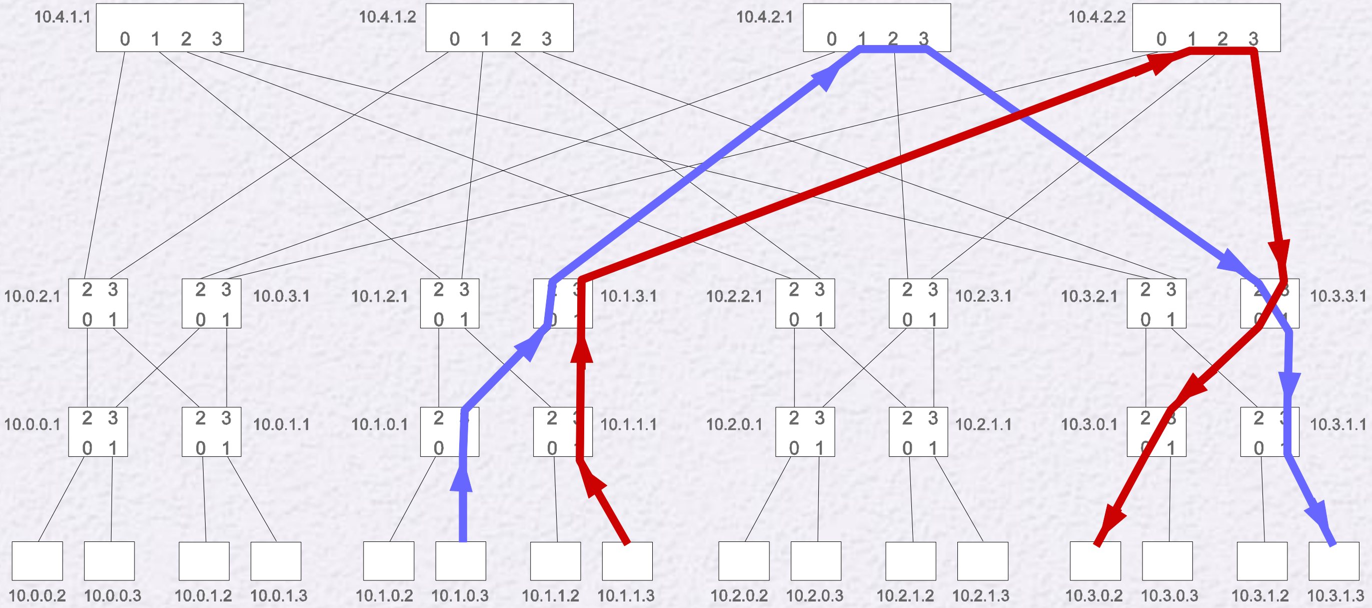
M. Al-Fares, A. Loukissas, A. Vahdat, “A Scalable, Commodity Data Center Network Architecture”, ACM SIGCOMM 2008

General Network Monitoring
Half-Consistent routing



Example of half-consistent routing in a fat-tree (based on IP address masks)

General Network Monitoring
Half-Consistent routing



Example of half-consistent routing in a fat-tree (based on IP address masks)

General Network Monitoring
Half-Consistent routing

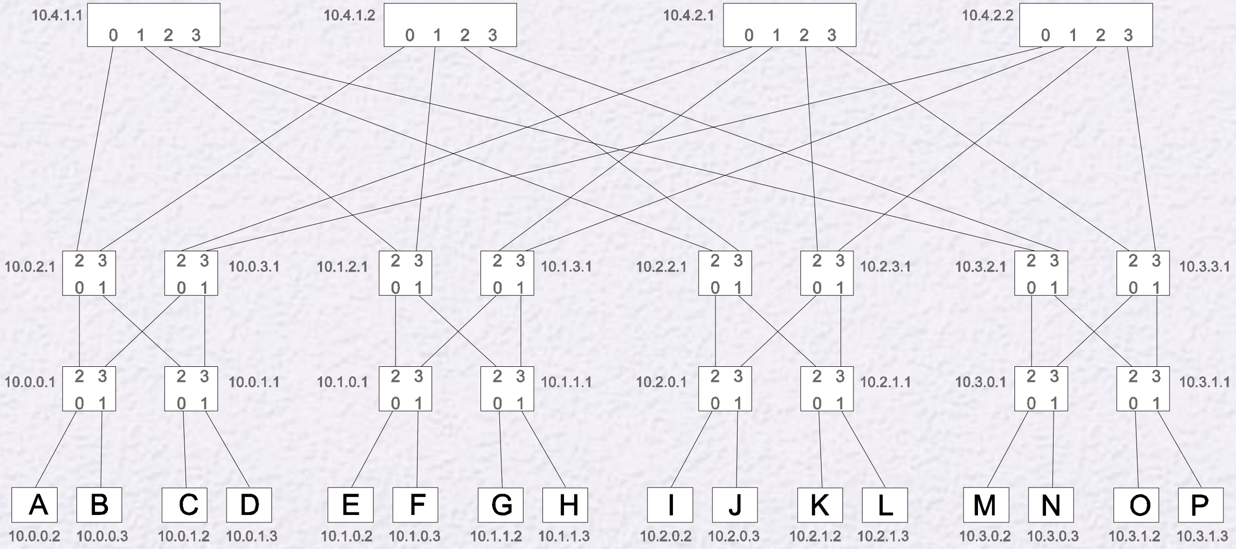
III. Half-consistent routing

Definition: *If a routing scheme guarantees that any path* $p_i \in P$ *can be divided into two segments* $s_1(p_i)$ *and* $s_2(p_i)$, *such that the property of routing consistency holds for the set* $P_{1/2} = \cup_{pi \in P} \{s_1(p_i), s_2(p_i)\}$, *then the routing scheme is called* half-consistent.

General Network Monitoring
Half-Consistent routing

Lemma:

Any shortest-path routing scheme on a fat-tree is half-consistent.



General Network Monitoring
Half-Consistent routing

Lemma. *Given a path $p_i \in P$ of maximum length d^*, under the assumption of half-consistent routing, with $m = |P| > 1$ monitoring paths, the maximum number of different encodings in the rows of $M(\hat{p}_i)$ is $\min\{2^{m-1}, 4 * (m-1), d^*\}$.*

General Network Monitoring
Half-Consistent routing

Theorem (Half-consistent routing)**.** *In a general network with n nodes, m > 1 monitoring paths, diameter d*, under half-consistent routing, the number of identifiable nodes is upper bounded by*

$$\psi^{hcr}(m,n,d^*) \leq \sum_{i=1}^{i_{\max}} \binom{m}{i} + \left\lfloor \frac{N_{\max} - \sum_{i=1}^{i_{\max}} i \cdot \binom{m}{i}}{i_{\max}+1} \right\rfloor$$

where $$i_{\max} = \max\{k \mid \sum_{i=1}^{k} i \cdot \binom{m}{i} \leq N_{\max}\}$$

and $N_{max} = m \cdot \min\{2^{(m-1)}, 4 \cdot (m-1); d^*\}$.

Performance evaluation

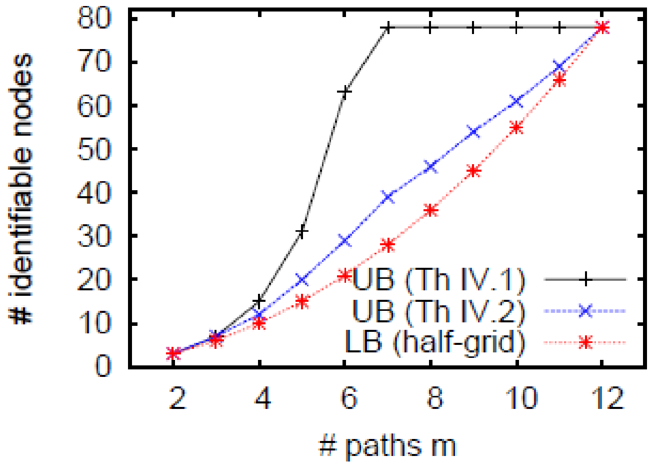


Fig. 14. Upper and lower bounds on the number of identifiable nodes in a half-grid network with $n = 78$, varying m, and $d^* = 12$.

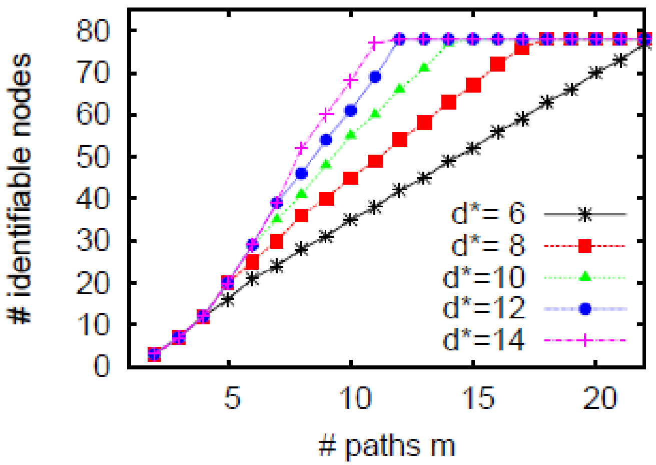


Fig. 15. Bound under consistent routing (Theorem IV.2) with varying number of paths and maximum path length (network as in Figure 14).

Performance evaluation

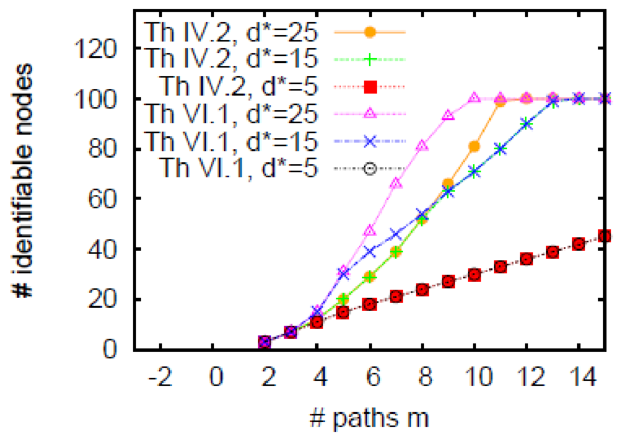


Fig. 20. Comparison between the bounds of Theorem IV.2 (consistent routing) and A.1 (half consistent routing) - 100 nodes, varying d^*.

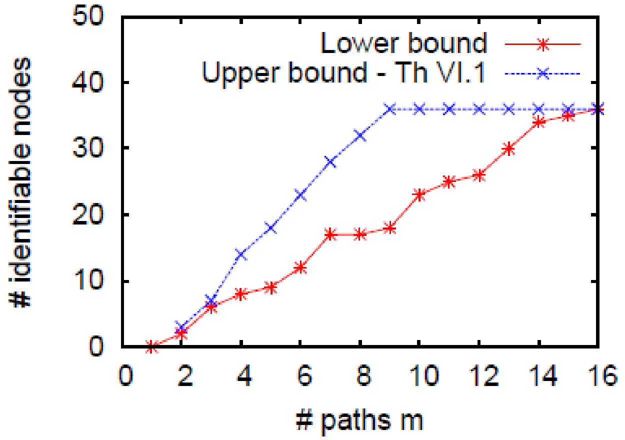


Fig. 21. Comparison between the bound of Theorem A.1 (half-consistent routing) and a lower bound for a 4-ary fat-tree with three layers.

Conclusions

- The problem of maximizing number of nodes whose states can be identified via Boolean tomography can be seen as graph-based group testing
- Upper bound on the number of identifiable nodes under different routing assumptions has been derived
- Provides insight for the design of topologies and monitoring schemes with high identifiability

Open problems

- Current bounds are topology agnostic. What if we know the adjacency matrix of our network topology?

- Algorithms for monitor deployment and path selection, with the objective to maximize node identifiability.

- We typically have partial knowledge and partial controllability.
  - Some nodes are known to be working, some others are known to be broken. There is a grey area where we want to assess damages. How does this change the algorithms?
  - Monitors can only be placed in our own routers. We don't own the entire network. What is the best we can do with the nodes that we can control?

- Some nodes/paths are more important than others, how can we design algorithms that prioritize identifiability of given nodes?

- Provide further insight for the design of topologies and monitoring schemes with high/low identifiability

Thank You!