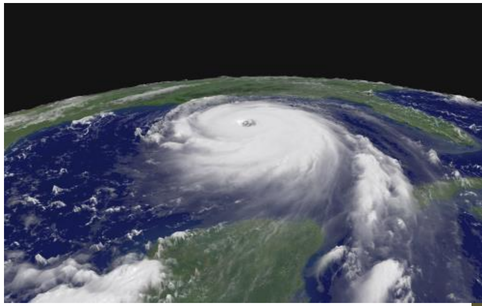
Network Recovery from Massive Failures under Uncertain Knowledge of Damages

Diman Zad Tootaghaj[1]
Hana Khamfroush[1]
Novella Bartolini[2]
Stefano Ciavarella[2]
Seamus Hayes[1]
Thomas La Porta[1]

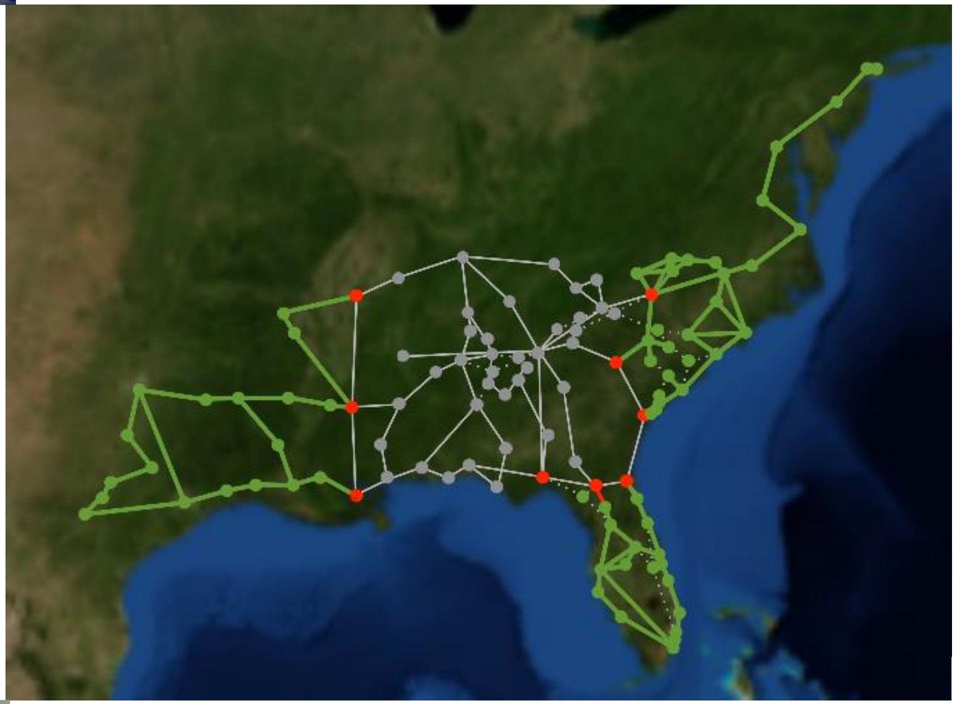
[1] The Pennsylvania State University (US)
[2] Sapienza University (Italy)

Outline

- Motivation
- Problem Definition
- Proposed Algorithms
- Evaluation
- Conclusion

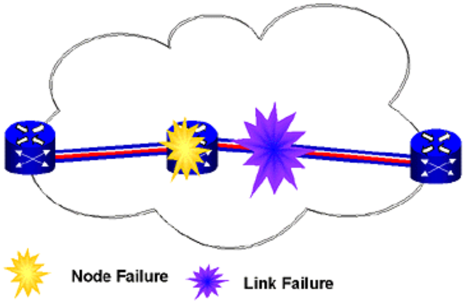


Natural disaster



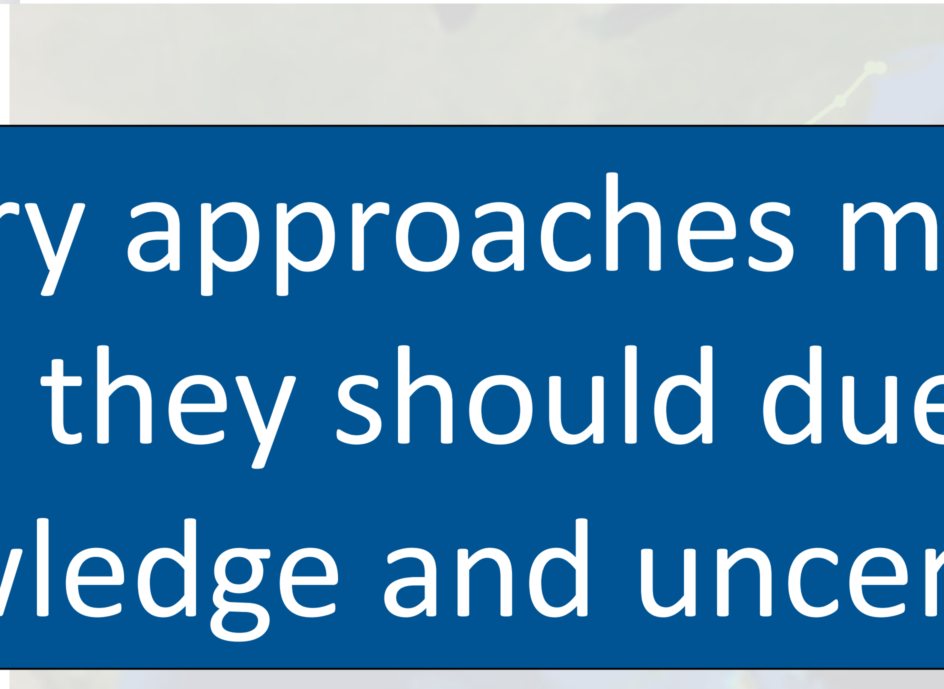
Malicious attacks

Hurricane

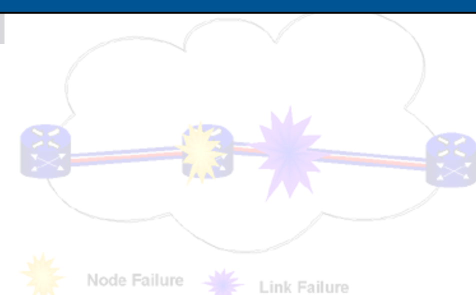


Random Failures

Natural



Recovery approaches may not work as they should due to lack of knowledge and uncertainty



Random Failures

Motivation

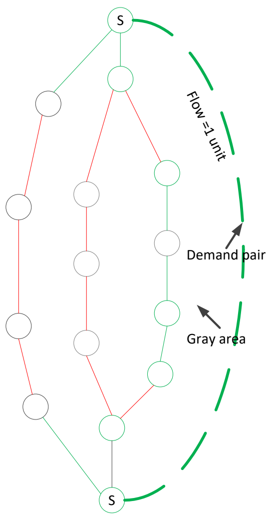
- **Problem**: Given a set of nodes and edges whose failure status is unknown, we would like to implement optimal recovery algorithm.
- **Observations**:
  - Large-scale **failures** in communication networks due to natural **disasters** or **malicious attacks** can severely affect critical communications and threaten lives of people.
  - In real-world scenarios, the failure pattern might be **unknown** or only **partially known**.

- **Key Idea**: Use multi-stage stochastic optimization to recourse as more information becomes available.
- **Results:**
  - Lower recovery cost (a factor of **3** on average).
  - Trade-off between:
    - Demand loss
    - Execution time
    - Repair cost

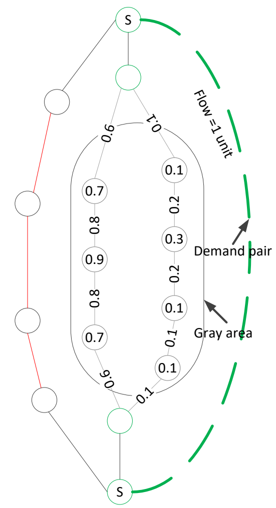
Motivation

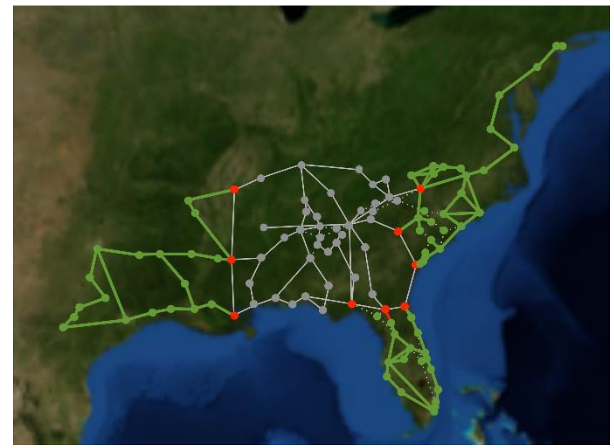
The failure pattern is unknown or only partially known

Full-knowledge



Partial-knowledge





Total Number of repairs in previous approaches and our proposed approach

Network Name	OPT full-info	ISP full-info	ISP uncertain-info	Progressive ISP
BellCanada	28	34.23	79	45.39
Deltacom	36.94	43.26	112	55.5
KDL	55.2	63.2	165.65	83.55

Problem Definition

- We formulate the minimum expected recovery (**MINER**) problem as a mixed integer linear programming and show that it is NP-Hard.

$$\underset{\delta_i^v, \delta_{i,j}^e}{\text{minimize}} E_\zeta \sum_{(i,j) \in E_U \cup E_B} k_{ij}^e(\zeta_{ij}^e(n))\zeta_{ij}^e(n)\delta_{ij}^e + \sum_{i \in V_U \cup V_B} k_i^v(\zeta_i^v(n))\zeta_i^v(n)\delta_i^v$$

$$\text{subject to} \quad c_{ij}.\delta_{ij}^v \geqslant \sum_{h=1}^{|E_H|} f_{ij}^h(n) + f_{ji}^h(n) \qquad \forall (i,j) \in E \tag{1a}$$

$$\delta_i^v.\eta_{max} \geqslant \sum\nolimits_{(i,j) \in E_B} \delta_{ij}^e \qquad \forall i \in V \tag{1b}$$

$$\sum\nolimits_{j \in V} f_{ij}^h(n) = \sum\nolimits_{k \in V} f_{ki}^h(n) + b_i^h(n) \qquad \forall (i,h) \in V \times E_H \tag{1c}$$

$$f_{ij}^h(n) \geqslant 0 \qquad \forall (i,j) \in E, h \in E_H \tag{1d}$$

$$\delta_i^v, \delta_{i,j}^e \in \{0,1\} \tag{1e}$$

- The nodes and edges in the graph belong to:
  1. the sets $E_B \in E$ and $V_B \in V$ are known to be failed,
  2. the sets $E_U \in E$ and $V_U \in V$ are in the gray area whose failure patterns is unknown,
  3. the sets $E_W \in E$ and $V_W \in V$ are in the green area which are known to be working correctly in the system.

Proposed Algorithms (ISR)

Key idea:

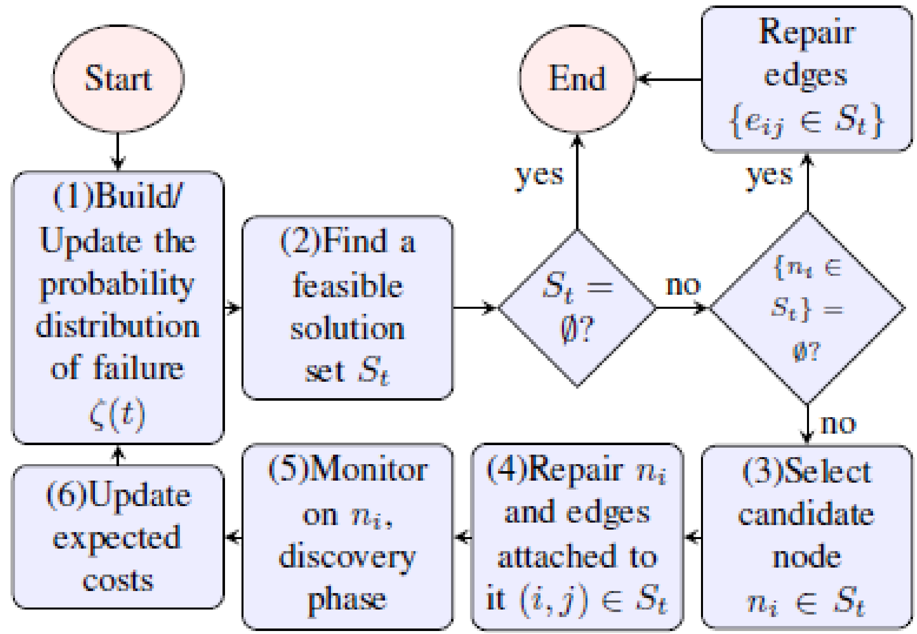
Initial guess (Failure probability distribution)

Find a feasible solution set (Not Optimal)

Select a candidate node to monitor repair and get more information

Update the initial belief

Recourse as more information becomes available.



Design principles:

Feasible Solution:

1. Iterative shortest path (ISR-SRT)
2. Iterative multicommodity (ISR-MULT)
3. Iterative branch and bound (ISR-BB)
4. Progressive ISP (P-ISP)

Best Candidate node selection: Betweenness centrality

$$N_i^* = argmax_{n_i \in S_t} \frac{\sum_{p \in P^*_{n_i}} f(p)}{\sum_{p \in P^*} f(p)}$$

Proposed Algorithms (ISR-SRT)

Key idea:

For each demand pair, finds the shortest paths.

Advantage:

- **Simple to implement,**
- **Polynomial time complexity.**

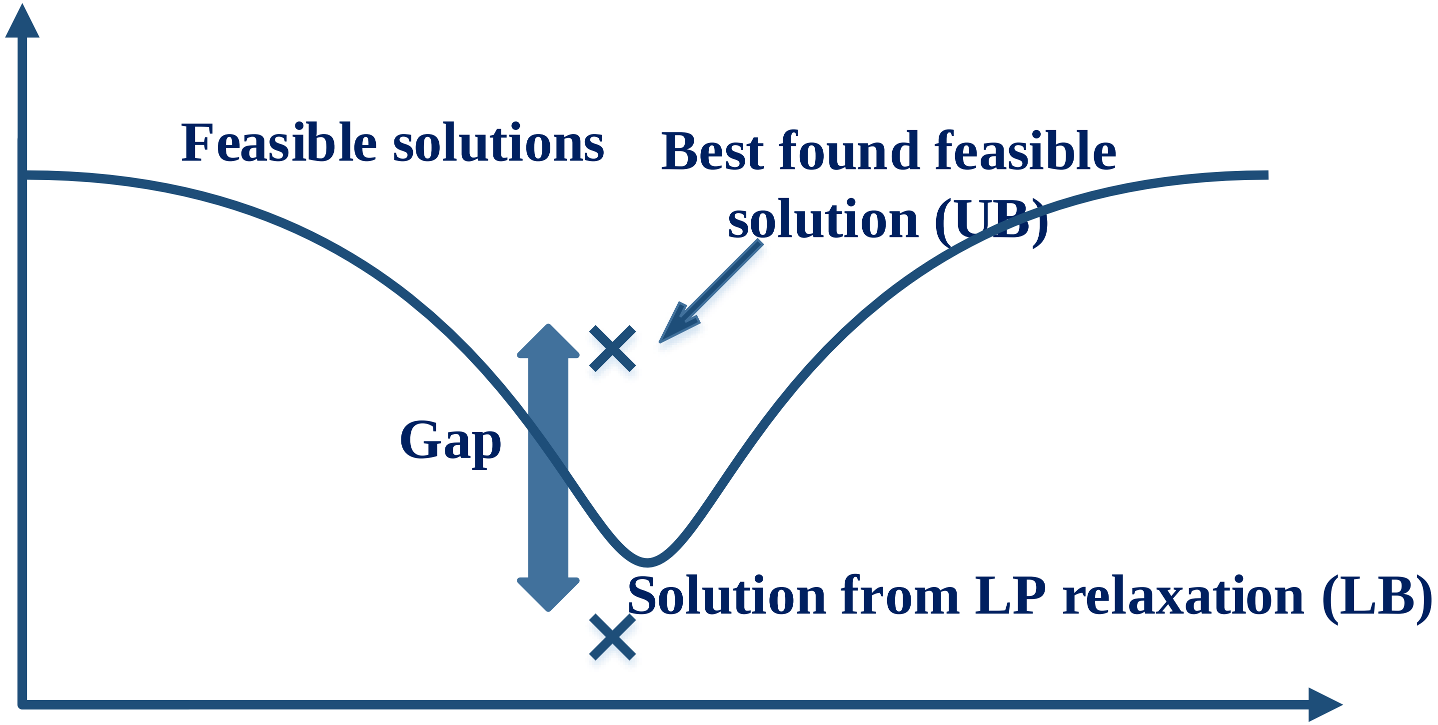
Problem:

- **Does not consider potential conflicts among demand pairs**
- **Demand Loss is possible.**

Proposed Algorithms (ISR-MULT)

Key idea:

LP relaxation of MINER and include all fractional variables in the current feasible set.

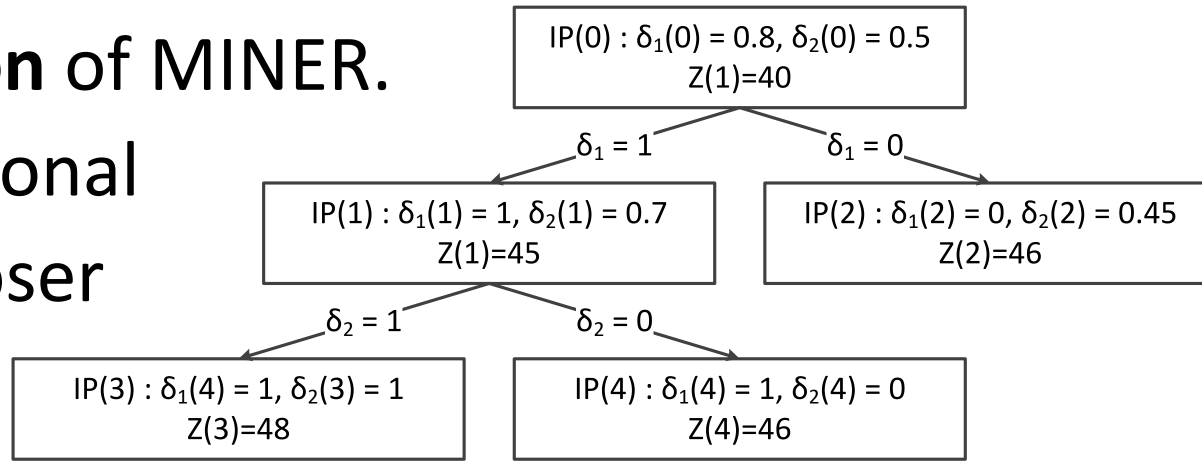


Proposed Algorithms (ISR-BB)

Key idea:

Solve **LP relaxation** of MINER.

Branch on a fractional variable to get closer to optimal.



Advantage:

- **Configurable trade-off between time complexity and optimality.**

Problem:

- **High execution time if runs for optimal solution.**

Proposed Algorithms (P-ISP)

Key idea:

We modify **Iterative split and prune** algorithm in [1].

- We use an uncertain estimation of failure distribution and change the length of the edge $e_{ij} \in E$ at the nth iteration to:

$$\frac{l^n(e_{ij})}{c_{ij}} = \frac{k_{ij}^e\left(\zeta_{ij}^e(n)\right)\zeta_{ij}^e(n) + \dfrac{k_i^v(\zeta_i^v(n))\zeta_i^v(n) + k_j^v\left(\zeta_j^v(n)\right)\zeta_j^v(n)}{2}}{c_{ij}}$$

Where $k_{ij}^e\left(\zeta_{ij}^e(n)\right)\zeta_{ij}^e(n)$, $k_i^v(\zeta_i^v(n))\zeta_i^v(n)$, and $k_j^v\left(\zeta_j^v(n)\right)\zeta_j^v(n)$ are the expected cost of repair for edge e_{ij} and nodes i and j based on the estimated probability distribution at the nth iteration.

[1] N. Bartolini et al. Network recovery after massive failures. In DSN 2016.

Outline

- Motivation
- Problem Definition
- Proposed Algorithms
- Evaluation
- Conclusion

Evaluation (Methodology)

Network Characteristics used in our evaluation:

- Real network from the Internet topology zoo [2].
- Synthetic Erdos-Renyi graphs with 100 nodes.

Network Name	# of nodes	# of edges	Average Node degree
BellCanada	48	64	2.62
Deltacom	113	161	2.85
KDL	754	895	2.37

Implementation:

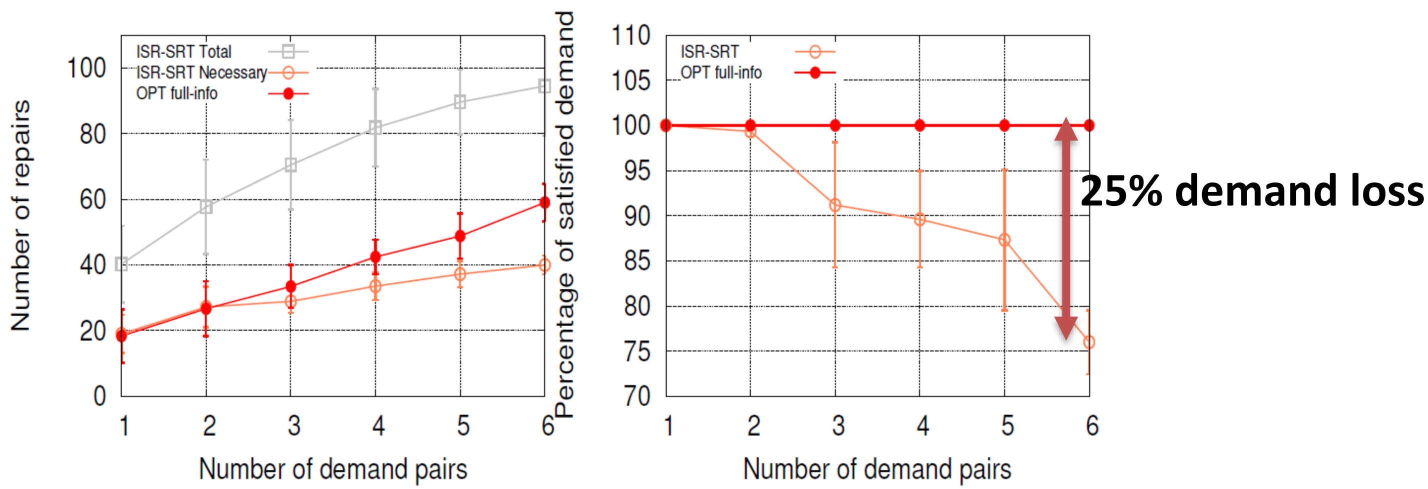
- **Python, Networkx, Gurobi optimization toolkit**

[2] The internet topology zoo. http://www.topology-zoo.org/, accessed in May, 2015.

Evaluation (demand loss)

Deltacom topology

- Trade-off between number of repairs and demand loss.



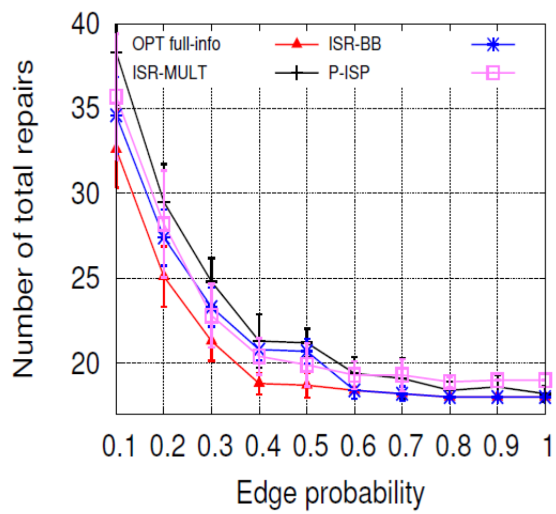
(a) repairs. (b) demand loss.

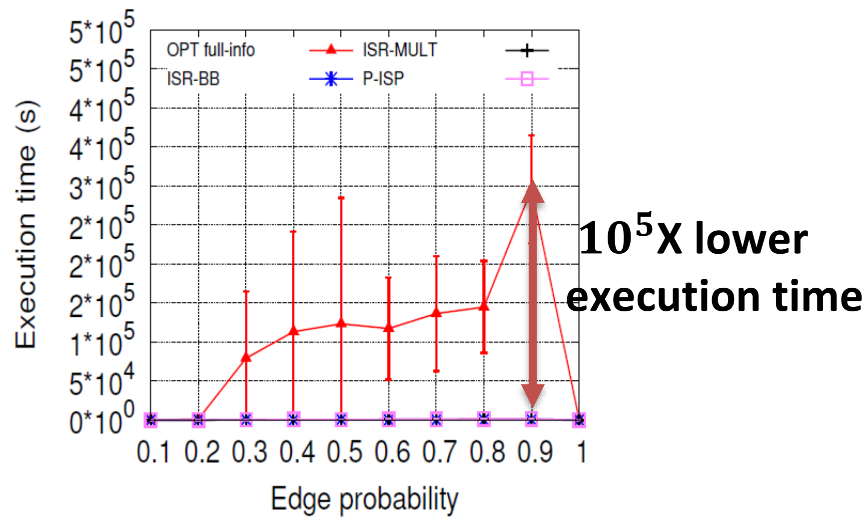
ISR-SRT has a low execution time while it does not consider potential conflict among demand pairs and has **25% demand loss** when the number of demand pairs is 6)

Evaluation (Execution time)

Synthetic Erdos-renyi (non-planar graph)

- Comparison of number of repairs and execution time.



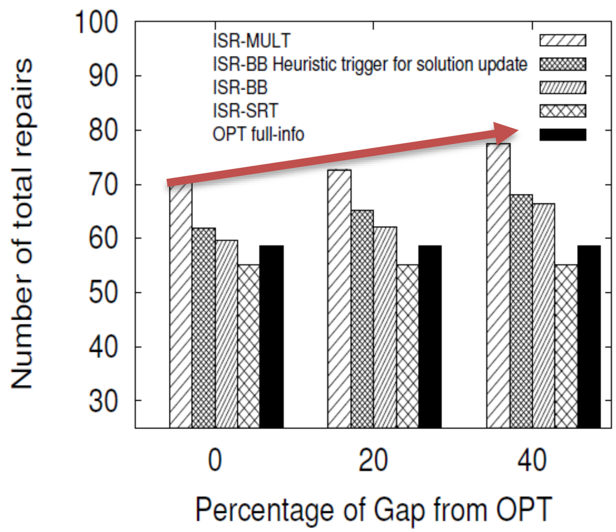


Solving the optimal NP-Hard problem can take 5 days to complete in some cases. Our ISR algorithm has **10^5X lower execution time** and it's recovery performance is close to optimal.

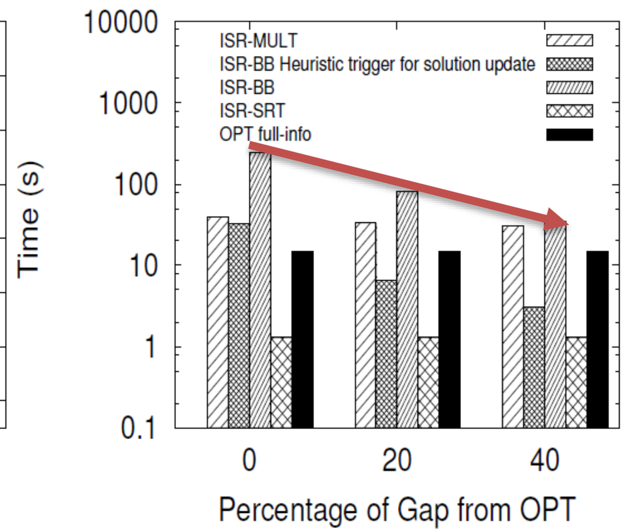
Evaluation (Trade-off)

DeltaCom topology

- Trade-off execution time and number of repairs.



(a) repairs.



(b) time.

Configurable trade-off

Using ISR-BB, we can have a configurable trade-off between the **number of repairs** and **execution time.**

Potential Implication of our work

Pos and Cons of each proposed approach

Algorithm	Cons	Pros
ISR-SRT	Demand loss, cannot satisfy all demands	Low complexity, easy to implement. Can be used to satisfy small critical demands in short time.
P-ISP	High number of unnecessary repairs in high demand load	Low time complexity compared to ISR-BB and ISR-MULT, works better than ISR-MULT in low demand load
ISR-BB	High time complexity due to large space exploration	Low number of repairs, best for small topologies. Can be configured to reduce the execution time with higher number of repairs
ISR-MULT	Moderate time complexity, high number of repairs in smaller traffics (can be combined with P-ISP to have advantage of both)	Smaller number of repairs compared to P-ISP, higher than ISR-BB. Better restoring performance for large number of demand flow/pair.

Conclusion

- **Problem**: Given a set of nodes and edges whose failure status is unknown, we would like to implement optimal recovery algorithm.
- **Observations**:
  - Large-scale **failures** in communication networks due to natural **disasters** or **malicious attacks** can severely affect critical communications and threaten lives of people.
  - In real-world scenarios, the failure pattern might be **unknown** or only **partially known**.

- **Key Idea**: Use multi-stage stochastic optimization to recourse as more information becomes available.
- **Results:**
  - Lower recovery cost (a factor of **3** on average).
  - Trade-off between:
    - Demand loss
    - Execution time
    - Repair cost

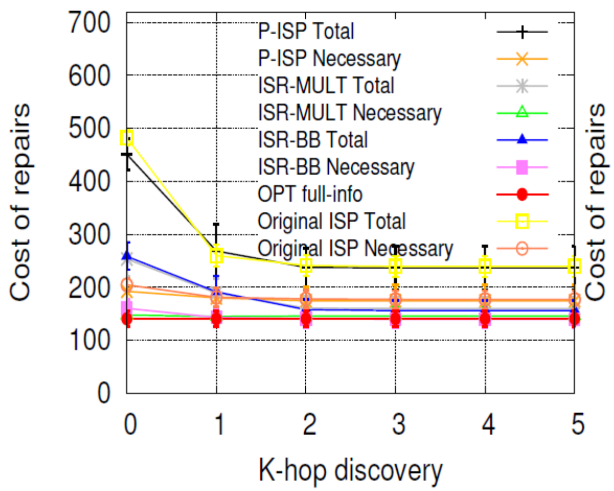
Questions?

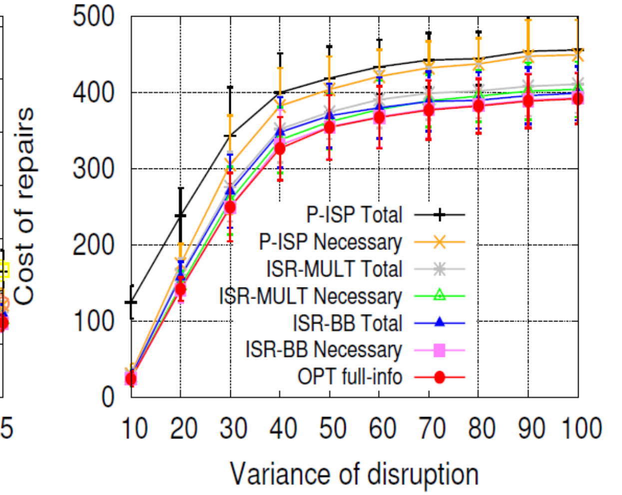
Hidden Slides

Evaluation (discovery, disruption)

BellCanada topology

- Discovery and disruption variation impact.



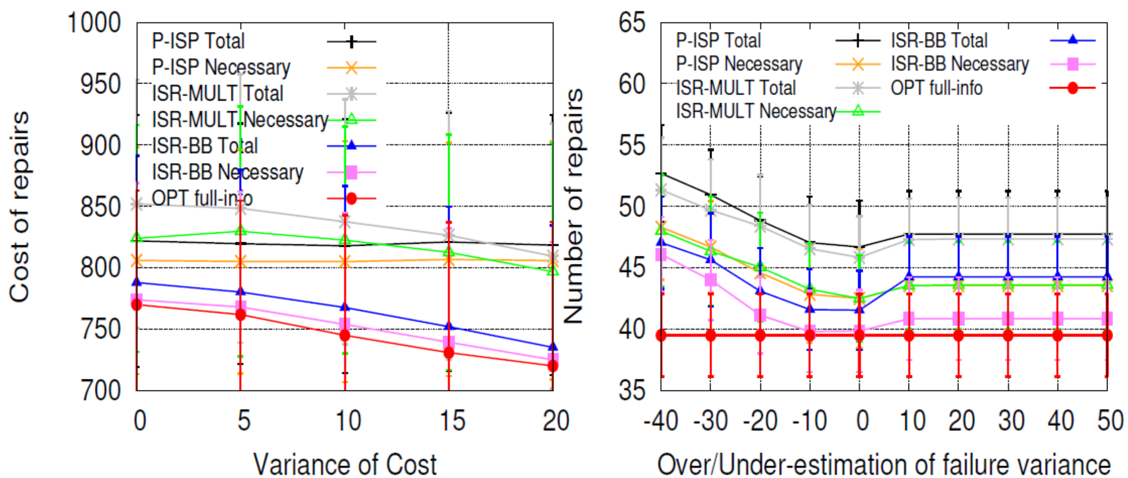


Increasing the number of discovered hops improves the restoring performance. As we increase the disruption variation, the total number of repairs increases until the whole network gets disrupted.

Evaluation (heterogeneous cost, sensitivity analysis)

<u>BellCanada topology</u>

- Heterogeneous repair cost and Sensitivity analysis.

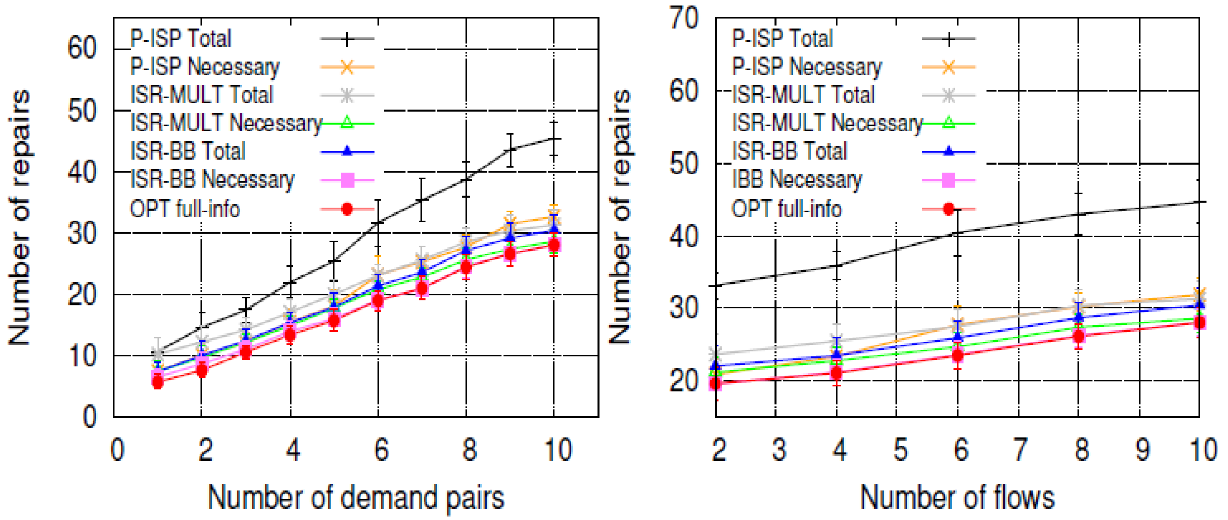


Our recovery approaches perform better when the variance of heterogeneity is higher. Underestimating the disruption, lead to routing of the critical demands through a part of network, which is more likely to be failed and therefore the recovery performance is lower.

Evaluation (discovery, disruption)

<u>BellCanada topology</u>

- Discovery and disruption variation impact.



(a) Increasing demand. (b) Increasing flows.

Increasing the load (demand pair or number of flows) leads to higher number of repairs.