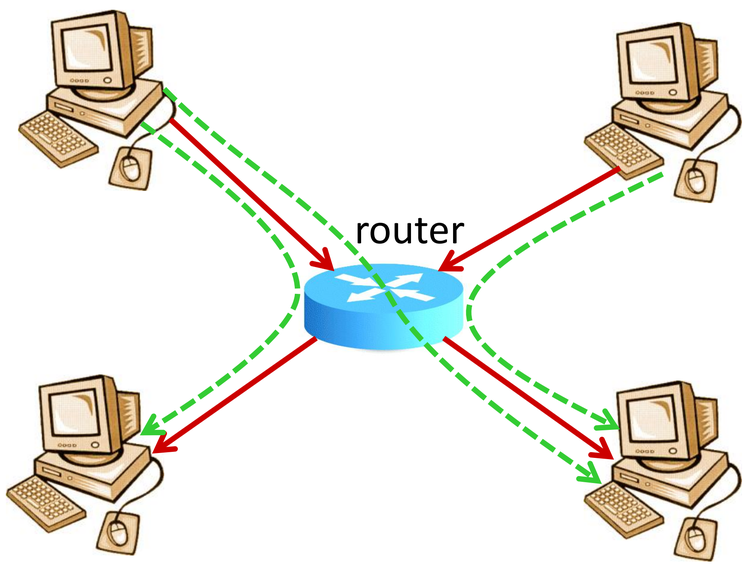
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences
Laboratory for Computer Communications and Applications & Operating Systems Lab

# The Battle for Non-Independence in Network Tomography

Denisa Ghita, Katerina Argyraki, Patrick Thiran

## 1. Network Tomography

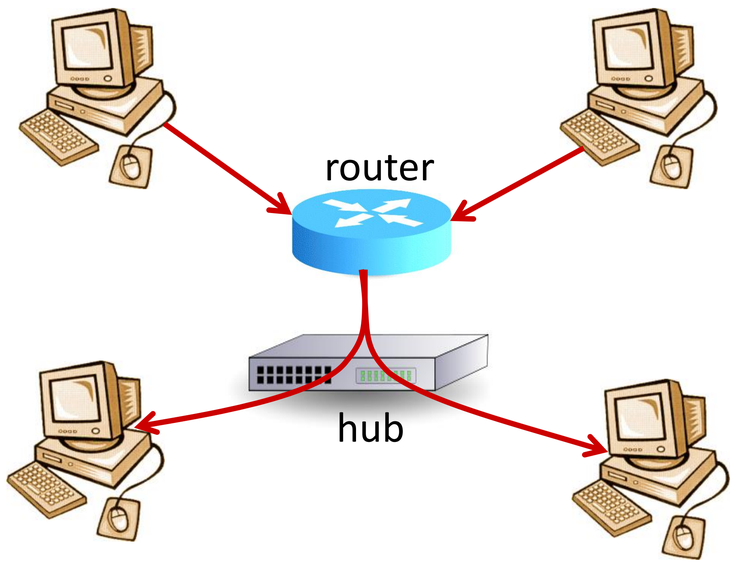
Network Tomography determines characteristics of *links* from *path* measurements.




Existing tomographic techniques work at IP level and assume *link independence*.

## 2. Link Correlation

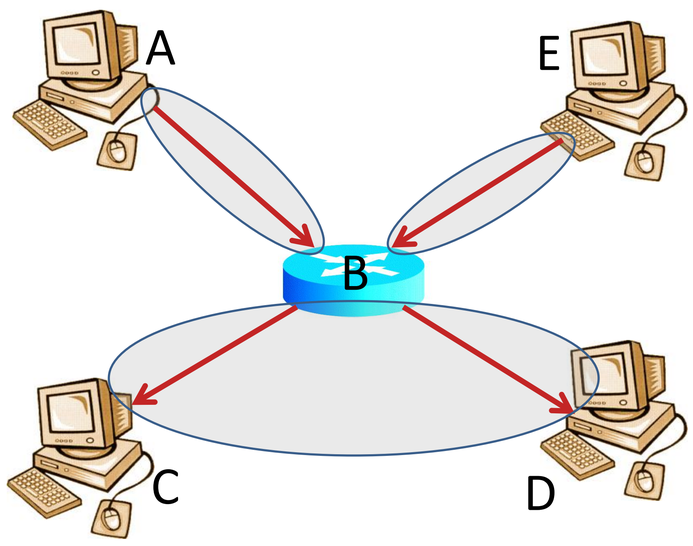
Two IP links might share a physical link.




If the router-hub link is faulty, both IP links that share it are faulty, hence *not independent.*

## 3. Correlation Model

Links are partitioned in correlation sets. Links in a set might be correlated, but are independent from links in other sets.

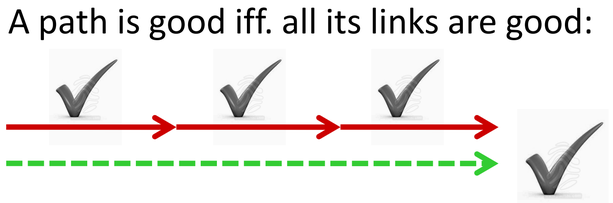



Links $e_{BC}$ and $e_{BD}$ belong to the same correlation set, they are independent from links $e_{AB}$ and $e_{EB}$.

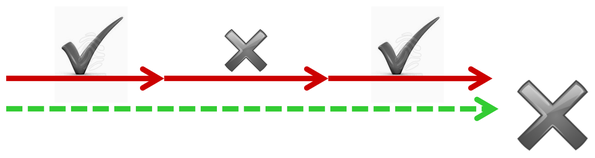
## 4. Boolean Tomography

A link is either good ✓ or faulty ✗.

A path is good iff. all its links are good:

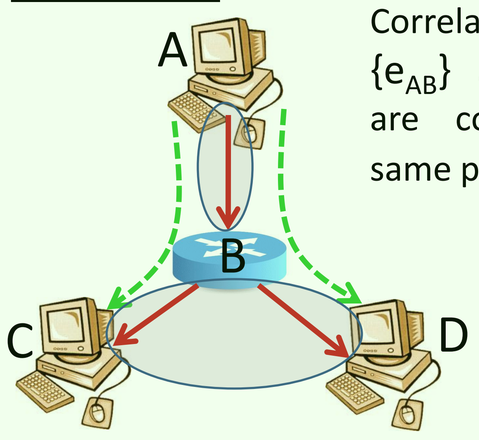


A path is faulty iff. at least one of its links is faulty:



*If links are correlated, can we find the probability that each link is faulty from the probabilities that paths are faulty?*

## 6. The Gist



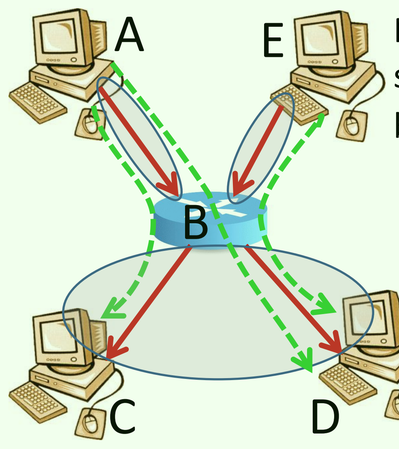

Correlation subsets $\{e_{AB}\}$ and $\{e_{BC}, e_{BD}\}$ are covered by the same paths: $P_{AC}, P_{AD}$

$\mathbf{P}(P_{AC}\text{ good}) = \mathbf{P}(e_{AB}\text{ good})\,\mathbf{P}(e_{BC}\text{ good})$

$\mathbf{P}(P_{AD}\text{ good}) = \mathbf{P}(e_{AB}\text{ good})\,\mathbf{P}(e_{BD}\text{ good})$

$\mathbf{P}(P_{AC}, P_{AD}\text{good}) = \mathbf{P}(e_{AB}\text{good})\,\mathbf{P}(e_{BD}, e_{BC}\text{ good}) \neq \mathbf{P}(e_{BD}\text{good})\mathbf{P}(e_{BC}\text{ good})$

~~Solvable~~: *4 unknowns and 3 equations!*




No two correlation subsets are covered by the same paths.

$\mathbf{P}(P_{AC}\text{ good}) = \mathbf{P}(e_{AB}\text{ good})\,\mathbf{P}(e_{BC}\text{ good})$

$\mathbf{P}(P_{AD}\text{ good}) = \mathbf{P}(e_{AB}\text{ good})\,\mathbf{P}(e_{BD}\text{ good})$

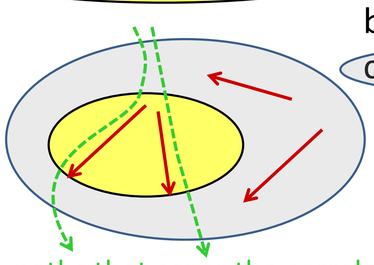
$\mathbf{P}(P_{ED}\text{ good}) = \mathbf{P}(e_{EB}\text{ good})\,\mathbf{P}(e_{BD}\text{ good})$

$\mathbf{P}(P_{AD}, P_{ED}\text{good}) = \mathbf{P}(e_{AB}\text{good})\,\mathbf{P}(e_{EB}\text{ good})\mathbf{P}(e_{BD}\text{good})$

*Solvable: 4 unknowns and 4 equations!*

## 5. Theoretical Results

correlation subset = set of links that belong to the same correlation set.

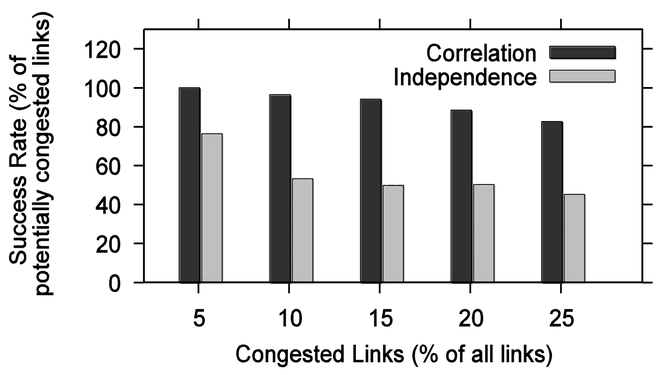


paths that cover the correlation subset

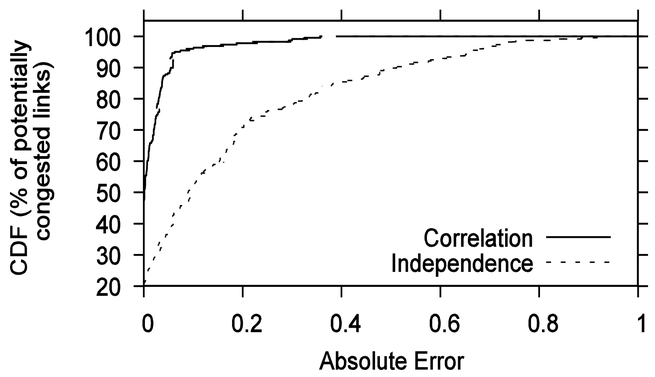
**Correlated Links Theorem [1]:**
*The probabilities of links being faulty are identifiable if each correlation subset is covered by a different set of paths.*

**Independent Links Theorem [2]:**
*The probabilities of links being faulty are identifiable iff. each link is covered by a different set of paths .*

## 7. Evaluation




Succes Rate= % of links with an absolute error below 0.1




CDF of absolute errors when 10% of links are faulty

## 8. Conclusion

We study network tomography on *correlated links.*

We *formally prove* under which conditions the probabilities of links being faulty are identifiable.

Our *tomographic algorithm* determines accurately the probabilities that links are faulty in a variety of congestion scenarios.

denisa.ghita@epfl.ch

[1] Network Tomography on Correlated Links, D. Ghita and K. Argyraki and P. Thiran, *submitted to IMC 2010*
[2] The Boolean Solution to the Congested IP Link Location Problem: Theory and Practice, H. Nguyen and P. Thiran, *Infocom 2007*