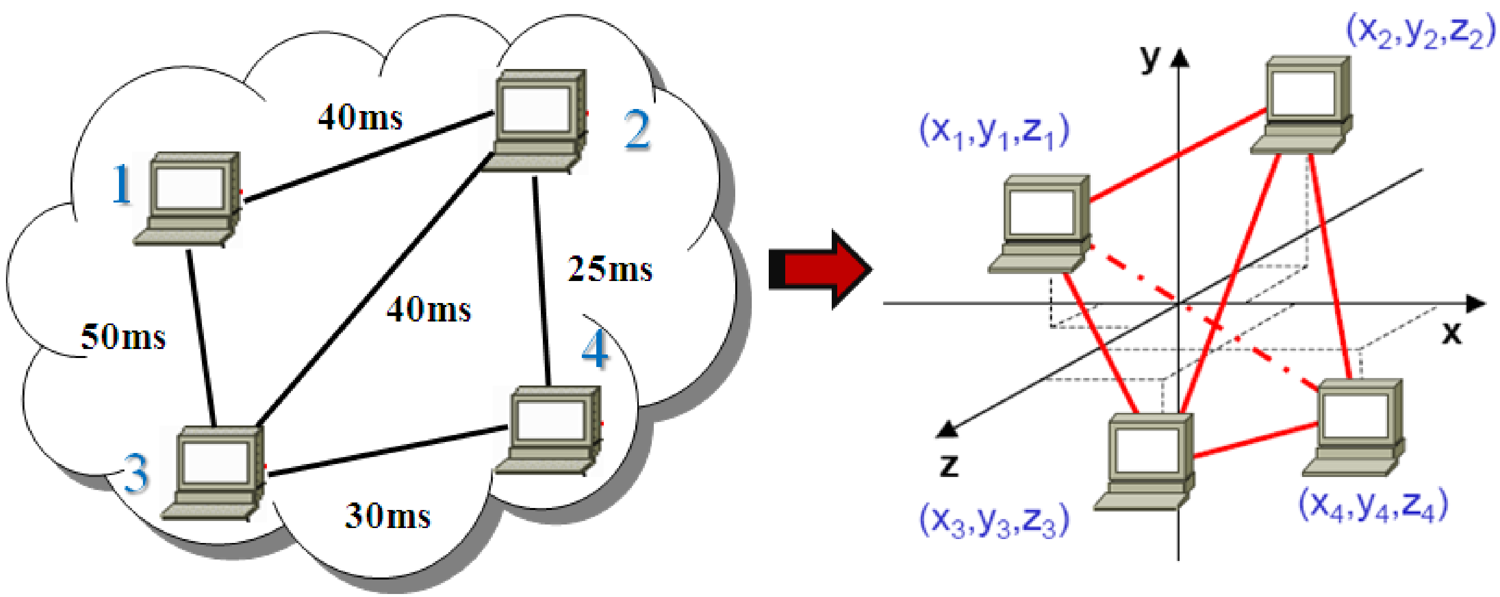
ROUTE OPTIMIZATION BASED ON THE DETECTION OF TRIANGLE INEQUALITY VIOLATIONS

Papa Ousmane Sangharé
Joint work with Bamba Gueye and Ibrahima Niang
University Cheikh Anta Diop of Dakar
papa.sanghare@ucad.edu.sn

Cape Town, South Africa
12–14 December 2011

Network Coordinate Systems (NCS)

- Model the Internet in a geometric space
- Associate with each node coordinates in this geometric space
- Estimate the network distance (RTT) between two pairs of nodes based on their coordinates



Network Coordinate Systems (NCS)

- Goals
  - Estimate de time beetween pairs of nodes
  - Limit the load of traffic related to direct measurements

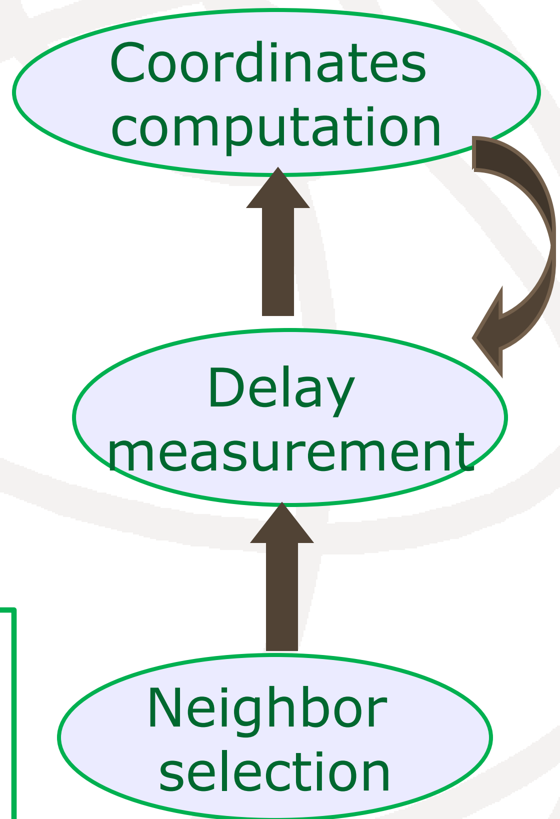
- Motivations

Useful for applications using the notion of proximity network

  - Peer to peer file sharing applications
  - Online game applications
  - Selection services of closest server

Network Coordinate Systems (NCS)

- Mechanism for calculating coordinates [imc08]



Centralized Coordinate Systems

Decentralized Coordinate Systems

Network Coordinate Systems (NCS)

- **Centralized coordinate systems**
- These systems involve a central component called Landmarks or reference nodes
- The nodes perform these measurements with landmarks to compute their own coordinates

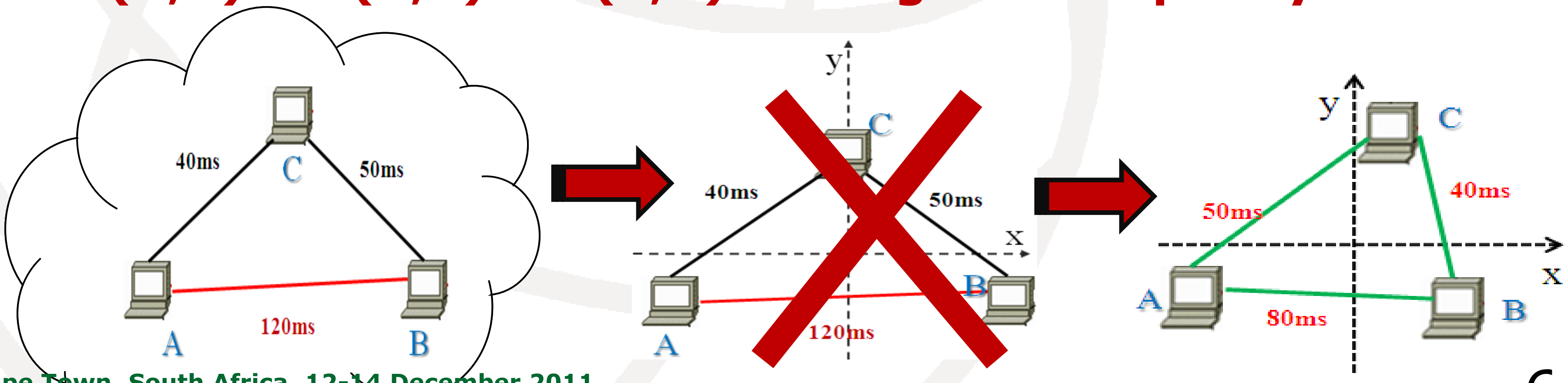
- **Decentralized coordinate systems**
- Generalize the role of landmarks to any node existing in the system, or by eliminating the landmark infrastructure
- The nodes perform these measurements with any node in the system to compute their own coordinates

Network Coordinate Systems and Triangle Inequality violation

- Network Coordinate Systems often assume triangle inequality holds for the Internet.
- The triangle inequality is often violated due to the Internet's structure and routing policies (potatoe routing, path inflation).
- **Drawbacks** :

  Inaccuracies with respect to the estimation of the Network Coordinate Systems

d(A,B) > d(A,C) + d(C,B) Triangular Inequality Violation



Metrics for detecting TIV

- Prediction ratio [IMC 07] : defined by the following relation :*estimated distance / the measured (actual) distance*
  - **Drawback** **:**

    Do not take into account the heterogeneous selection of neighbors as suggested in vivaldi [SIGCOMM 04]

  - OREE [Networking 09] :

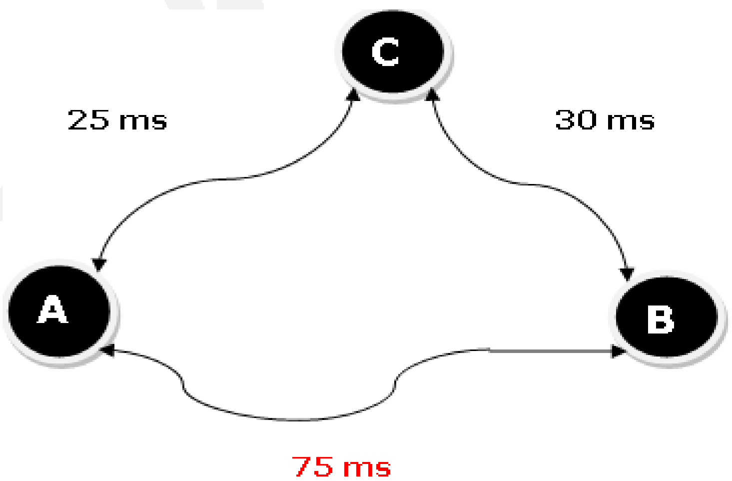
$$\text{OREE} = \frac{\textit{variance (estimated distances) – the measured distance}}{\textit{Average (estimated distances)}}$$

  - **Drawback** **:**

    OREE Uses a lot of information for detecting TIVs.

Triangle Inequality violation and the overlay Routing

- However TIVs may be also be useful in the routing overlay
  - By exploiting the detour paths they offer
- The TIV-base routing approach are benefit to:
  - Peer-to-peer applications
  - Online games
  - Distributed applications
  - VoIP



d(A,B) > d(A,C) + d(C,B).

Contributions

- Propose an effective metric for detection of TIVs
- Optimize routing in overlays network through TIVs
  - Detect the potential nodes that can act as shorcut for a given path using a clustering approach and new metric

Experimental setup

- P2psim as simulator
  - Implements a version of Vivaldi

- Three matrices delays (RTT) as dataset
  - P2Psim King dataset (1740 nodes)
    - 23% of links are TIV-Bases
  - Meridian dataset (2500 nodes)
    - 42% of links are TIV-Bases
  - PlanetLab dataset (180 nodes)
    - 9% of links are TIV-Bases

The new metric for detecting TIVs (RPMO)

- RPMO (Ratio of Prediction on Average Oscillations)metric considrs the following parameters
  - The Average oscillations distance between two nodes
  - The estimated distance
  - The actual distance

$$\boldsymbol{RPMO} = \frac{\textit{Estimated distance}}{\textit{RTT}} \times \frac{1}{\textit{Average oscillations}}$$

- RPMO keeps the heterogeneous selection of neighbors according to vivaldi
- RPMO Uses less computation overhead compared to OREE

The TIVs characteristics

- Two criteria indicate the severity of TIVs
  - The absolute severity

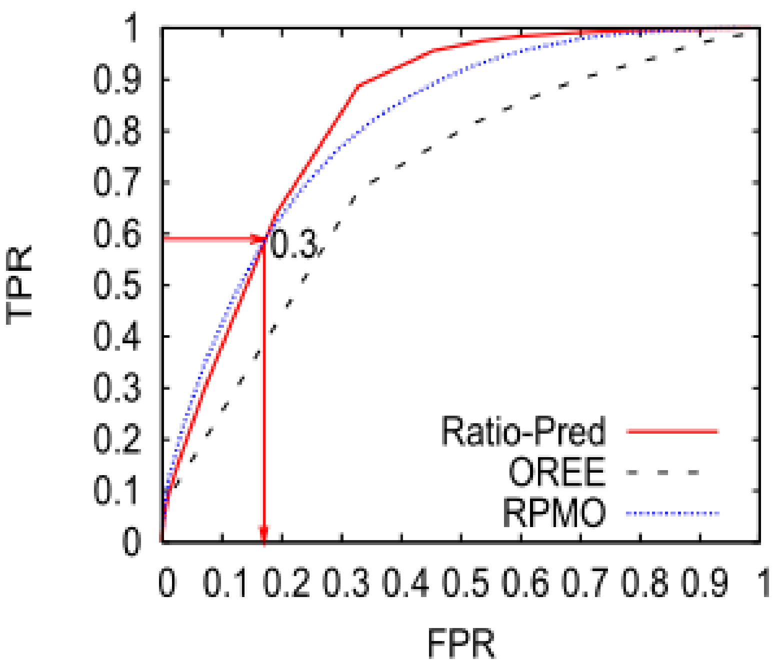
    $$Ga = d(A,B) - (d(A,C) + d(C,B))$$

  - The relative severity

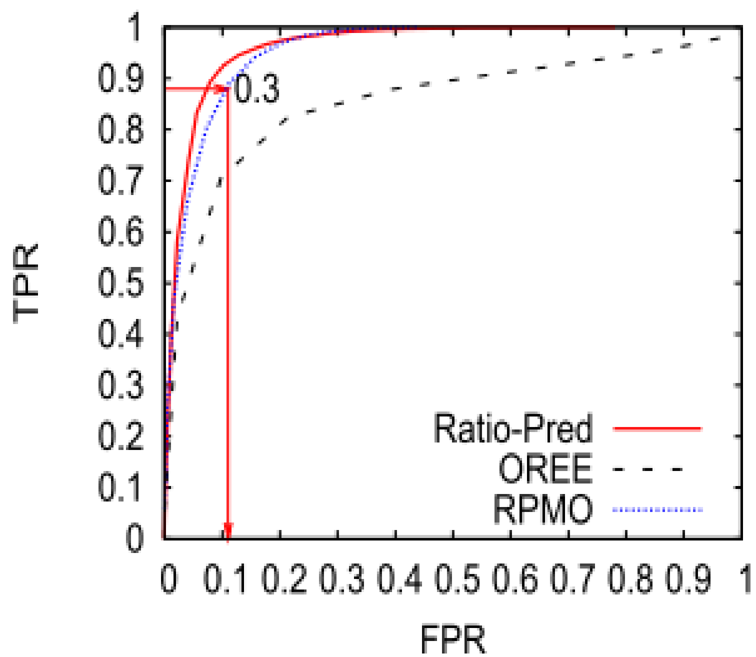
    $$Gr = \frac{d(A,B) - (d(A,C) + d(C,B))}{d(A,B)}$$

  - We are interested in TIVs thet meet Both creteria
    - **Ga > 10 ms and Gr > 0.1**

Comparison between RPMO, Prediction Ratio and OREE metrics



King dataset



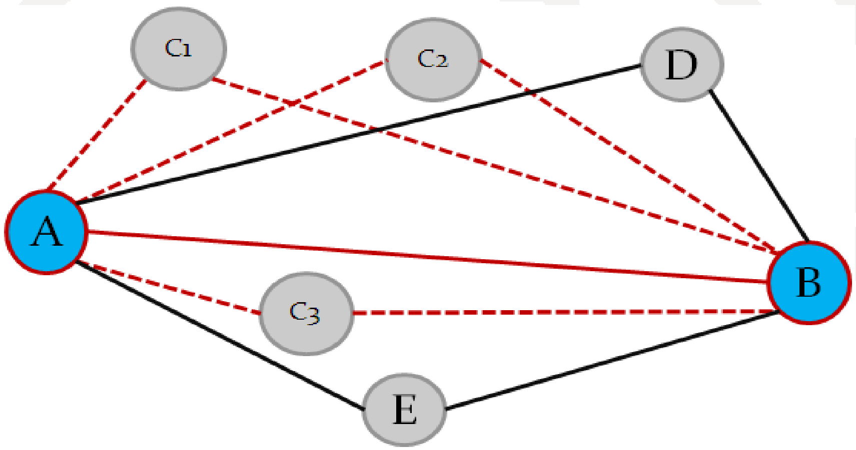
Meridian dataset

- RPMO is more efficient compared to OREE
- With the ratio of prediction the gap is reduced, we have roughly the same trend

OPTIMIZATION OF ROUTING IN THE OVERLAY NETWORK THROUGH TIVS DETECTION

- Goal

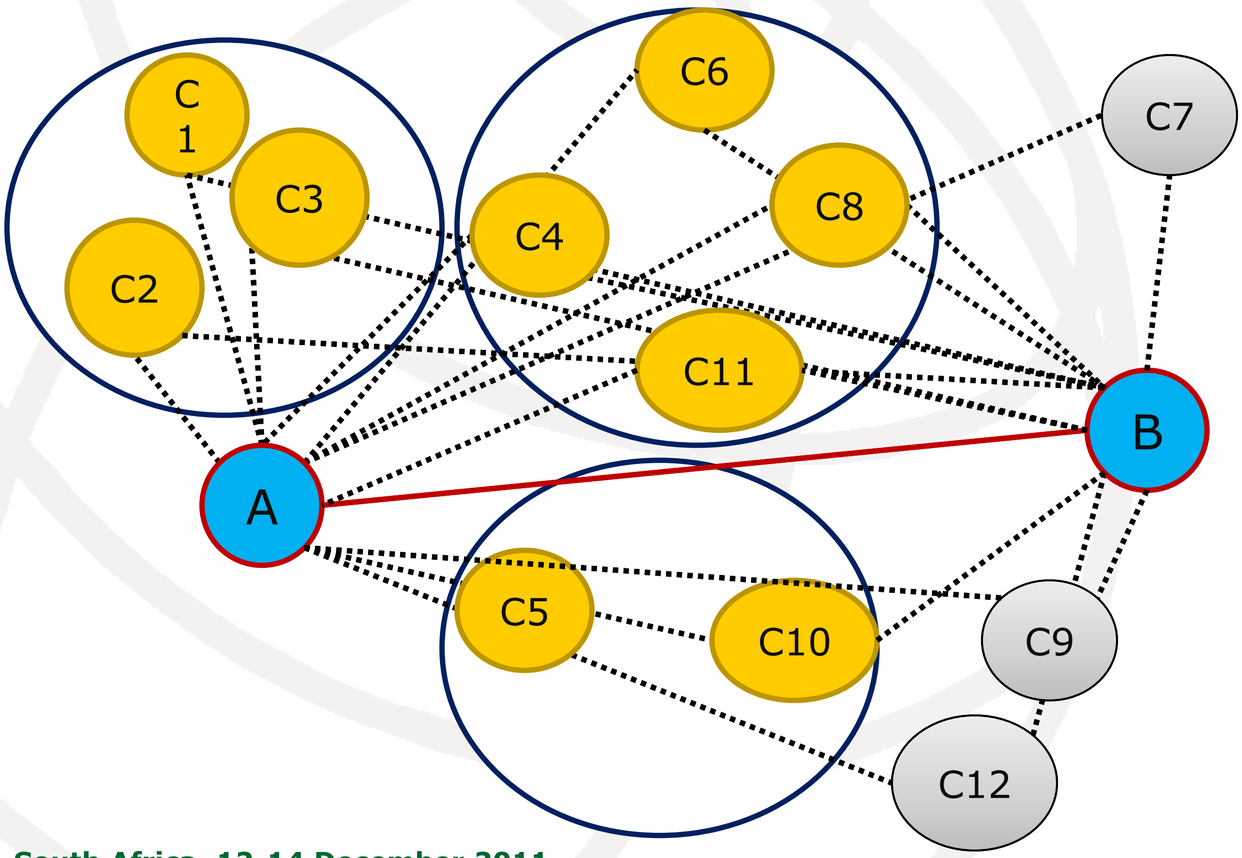
  for any link AB TIV-Base, find the potential point C, which provide the best detour paths



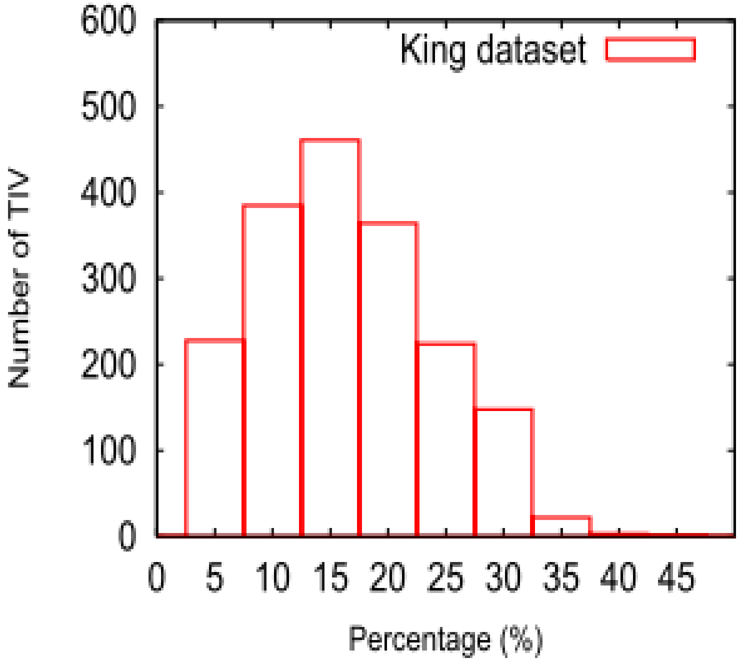
- Clustering method
- method of the detection metric of shortest paths

The clustering approach

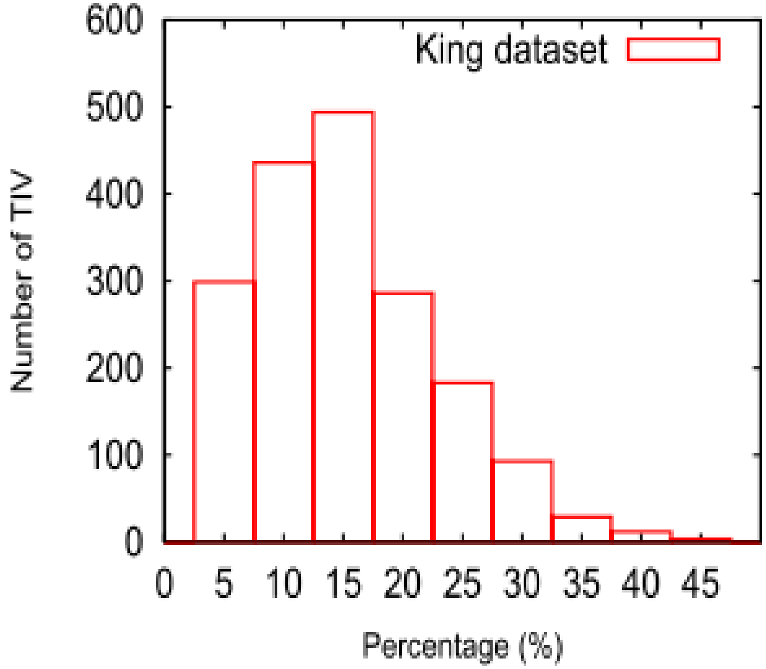
- Gather the potential points Ci of each link AB that is suspected as TIV-base



The clustering approach



Proportion of outliers (shortcuts) on the best shortcuts



Proportion of shortcuts of the"Best Cluster" on the best shortcuts

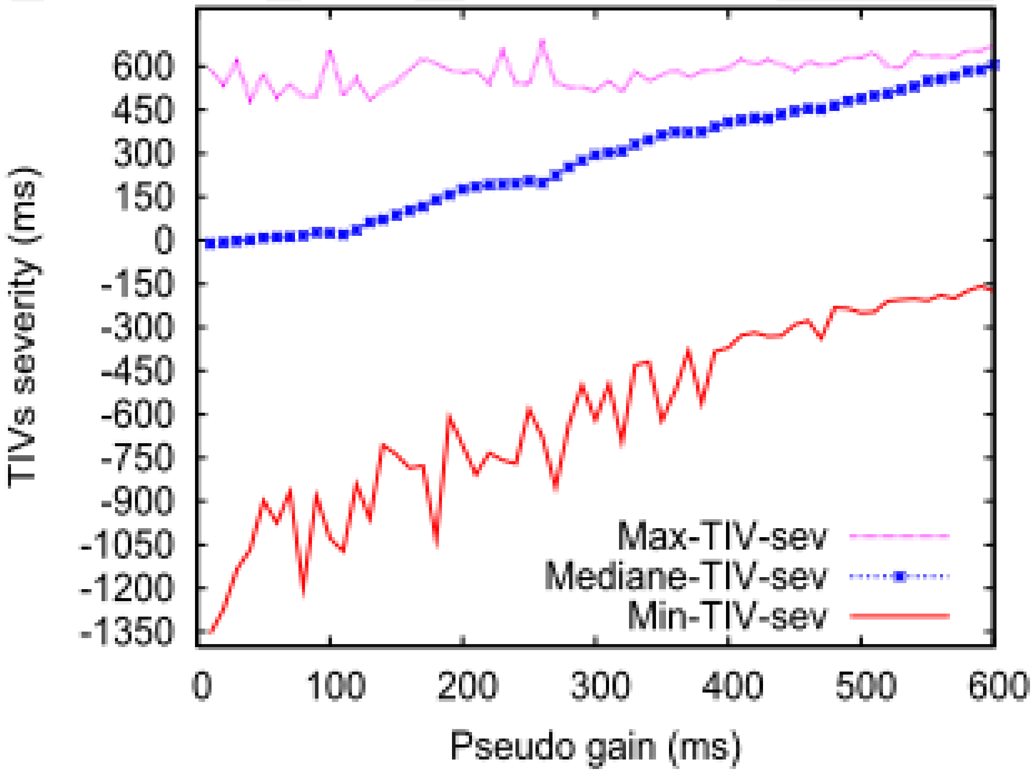
- The shortcuts detected by the both methods represent between 5 and 45% of the best shortcuts
- The clustering is not efficient for detecting the best shortcuts

MDGD (Metric for Detecting Good Detours) approach

- Relationship between pseudo gain and TIVs severity

  D'= d(AB) - (d (AC) + d (CB))

  Pseudo gain = d(AB) – (Estimée(AC)+Estimée(CB)).



- shows that more and more that the pseudo gain increases, we are dealing with triangles TIV-bases

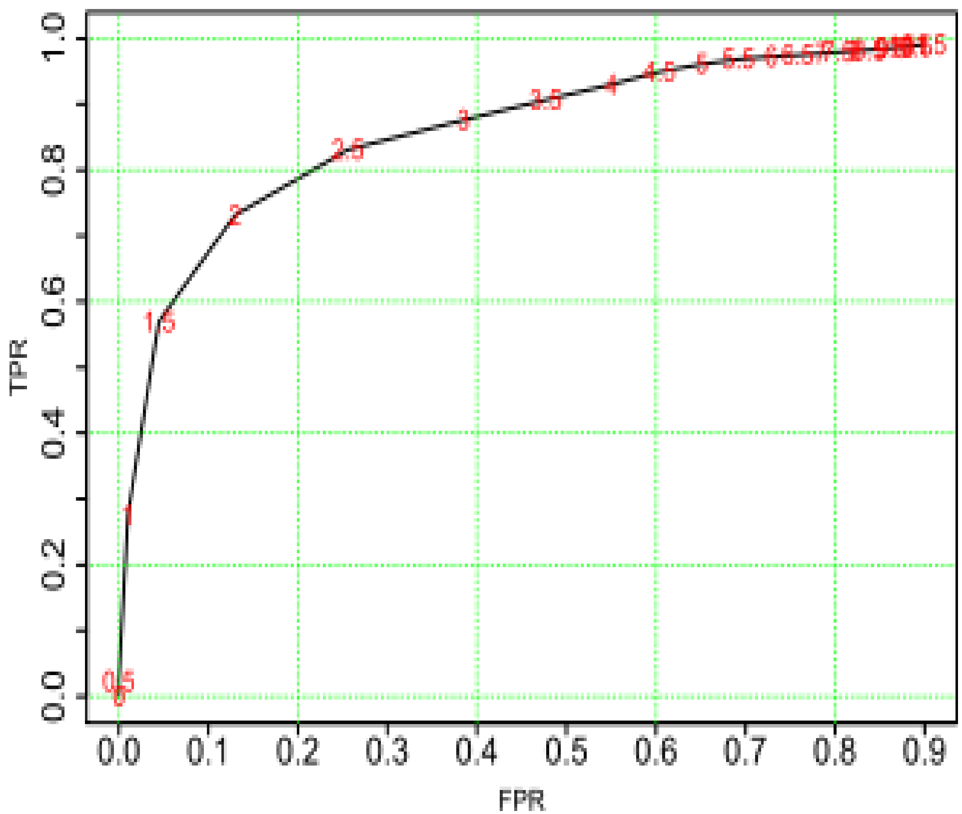
MDGD (Metric for Detecting Good Detours) approach

- Three parameters to build the MDGD metric:
  - The relative estimation error (**Er**) : d(AB)/Estimate(AB)
  - The absolute estimation error (**Ea**) : d(AB)−Estimate(AB)
  - Pseudo gain (**PG**) : d(AB)−(Estimate(AC)+Estimate(CB))

$$\mathbf{MDGD} = \frac{\mathbf{Er\ x\ Ea}}{\mathbf{PG}}$$

Evaluation of MDGD approach

- Evaluation of MDGD metric based on ROC curve



Threshold	TPR	FPR	Accuracy (ACC)
1.5	0.57	0.04	0.68
2	0.73	0.13	0.77
2.5	0.83	0.26	**0.81**
3	0.88	0.38	0.80

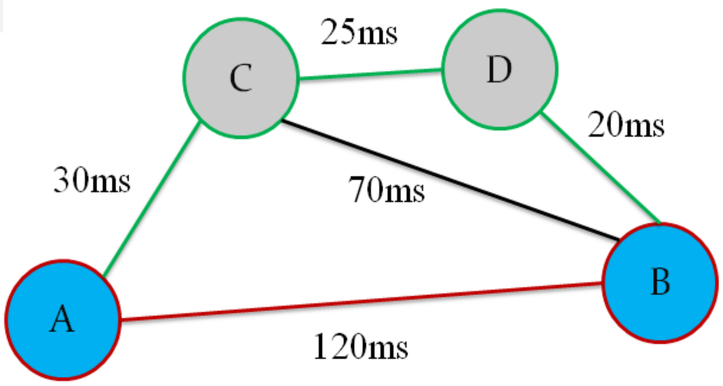
- 2.5 is best threshold that give us the best shortcuts with good accuracy

Conclusions

- Our metric RPMO outperforms OREE metric and presents the same trend with respect to the Prediction Ratio metric for the detection of the existence TIV.
- The use of the TIVs to improve routing in the overlays network (using the detour paths)
  - **MDGD** metric : remains an efficient solution

Perspectives

- Find a RPMO threshold value common for all datasets
- Research to shortcuts detours path with two jumps



d(A,C,D,B) < d(A,C,B) < d(A,B)

Thank you