SAC'22

Propagation Measure on Circulation Graphs for Tourism Behavior Analysis

Hugo Prevoteau[1,2], Sonia Djebali[2], Zhao Laiping[1], Nicolas Travers[2]

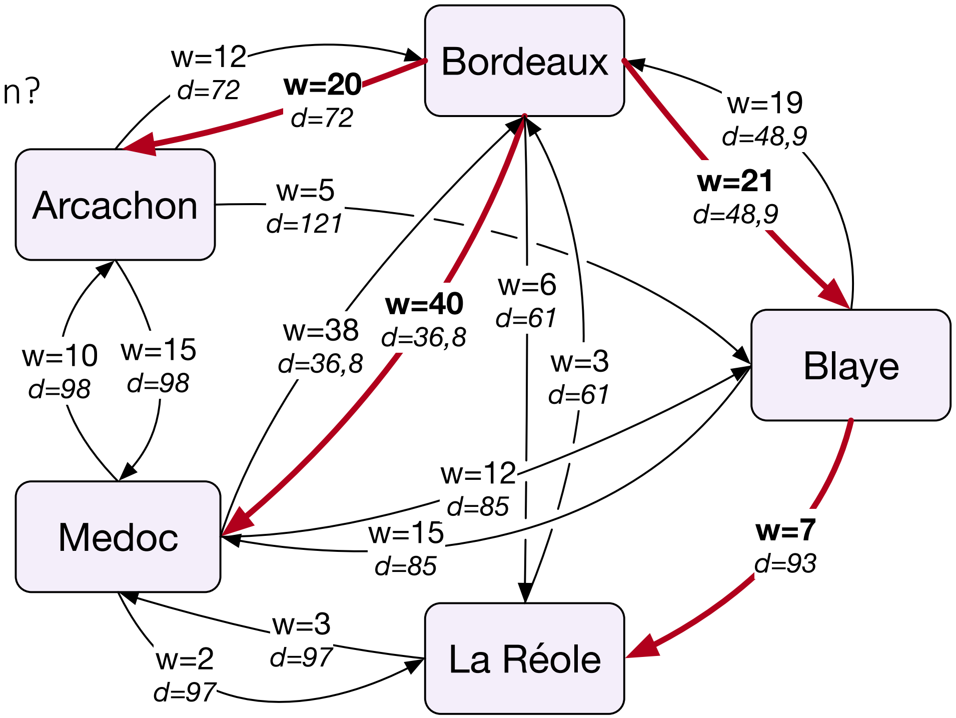
[1] College of Intelligence and Computing, Tianjin University, Tianjin, China
[2] De Vinci Research Center (DVRC), Paris La Défense, France

How to Analyze Tourists' Propagation?

Is it possible to use the Circulation Graph to understand propagation?
- Is there any logic?

Must take into account both topology & weights
- But also, we should take into account the distance

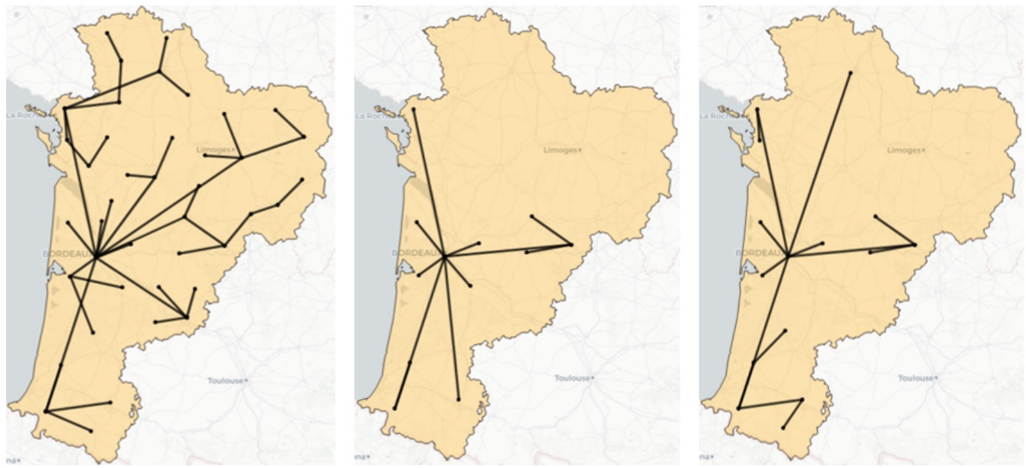
➢ Multi-weighted aggregated graphs



Minimum/Maximum Spanning Tree (MST)

➡️ Reflects the traffic flow and hierarchy in the underlying system [Stam et al. 2014]

MST: Kruskal algorithm [AMS'56]



French, American and Spanish Maximum Spanning Trees in 2018
Nouvelle-Aquitaine (district scale)

MST – How to compare both topologies and Remoteness?

- **Tree Edit Distance** - # of nodes/edges interchange
  - ➢ No viable comparison

- **Tree Hierarchy** - # of leafs vs links & Betweenness Centrality

$$T_h = \frac{L}{2 \times m \times BC_{max}}$$

  - ➢ Only dedicated to star vs lines topologies & no notions of distances.

Definition

Definition: Remoteness Influence Factor (RIF)

*Consider a Multi-weighted graph AC(V, E(w, d)) the RIF measures the **remoteness** of vertices combined with their **influences** in AC. For each node n ∈ V, it computes its normalized distance from a source* **s**, *combined with the inverse of its centrality BC(n). It is defined as:*

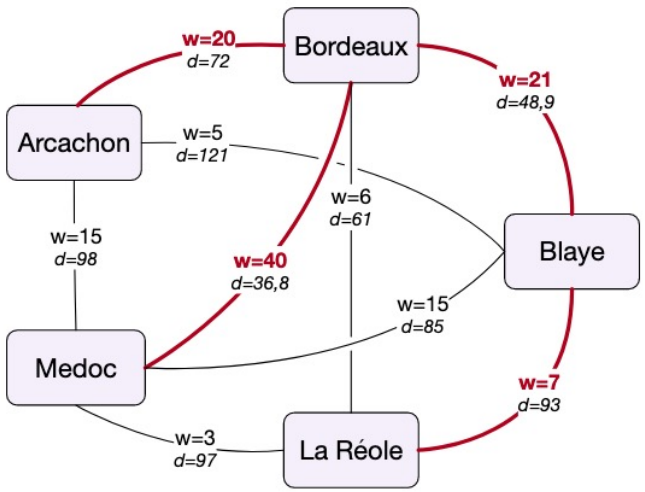
$$RIF(AC, s) = \sum_{n \in \mathcal{V}} \left(\frac{\log_{dist_{max}} \left(\sum_{e \in path(s,n)} dist(e) \right)}{|\mathcal{V}| - 1} \times \frac{\frac{1}{1+WBC(n)}}{|\mathcal{V}|} \right)$$

i.e., Max WBC(AC)

Normalized log(distance) from source **s**

Impact of nodes in AC

Remoteness Influence Factor - Example



$$RIF_{Fig1b} = \underbrace{\frac{log_{141.9}72}{4} \times \frac{\frac{1}{1+0}}{5}}_{Arcachon} + \underbrace{\frac{log_{141.9}36.8}{4} \times \frac{\frac{1}{1+0}}{5}}_{Médoc} + \underbrace{\frac{log_{141.9}48.4}{4} \times \frac{\frac{1}{1+0.5}}{5}}_{Blaye} + \underbrace{\frac{log_{141.9}141.9}{4} \times \frac{\frac{1}{1+0}}{5}}_{La\ Réole}$$

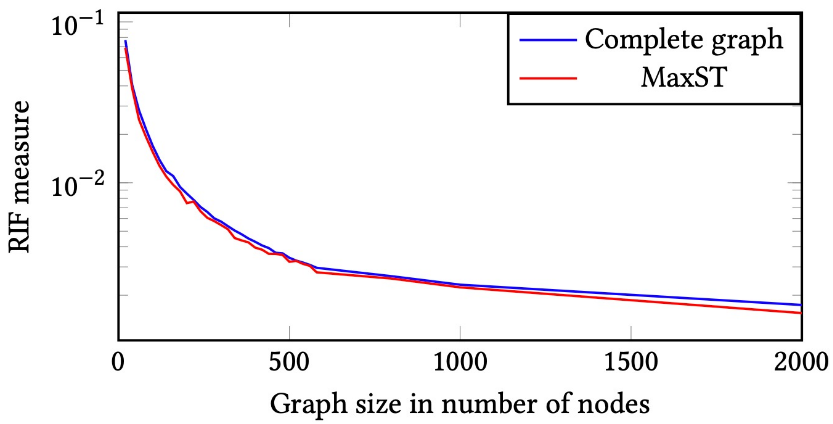
RIF Computation on a Graph

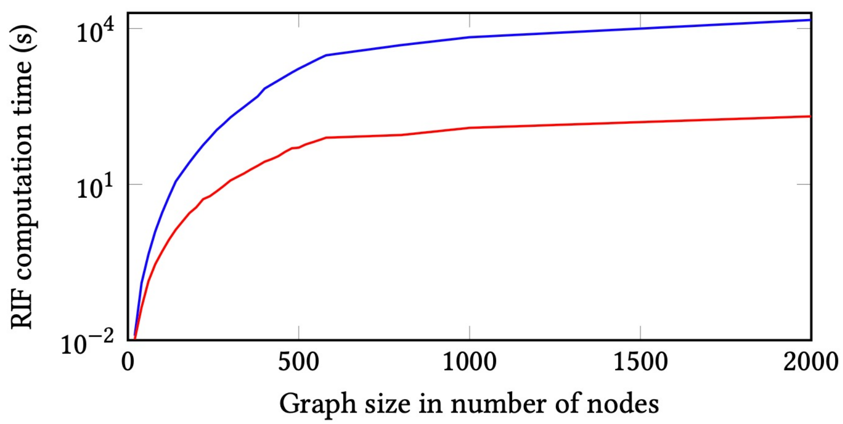
Algorithm 1 Computation of the Remoteness Influence Factor

Require: $\texttt{AC}(\mathcal{V}, \mathcal{E}(w, d))$ a graph, $s \in \mathcal{V}$ is the source node of the graph

function REMOTENESS(AC, s)
  wBetweennessCentrality = $\boxed{\texttt{WeightedBetweennessCentrality}(AC(w))}$
  distancePairs = $\texttt{Dijkstra}(AC(d), s)$
  max_dist = max(distancePairs)
  for $n \in \mathcal{V} - s$ **do**
    $rif = rif + \log_{max\_dist}(distancePairs[s][n]) \times \frac{1}{1+wBetweennessCentrality[n]}$
  end for
  return $\frac{rif}{(|\mathcal{V}|-1)\times|\mathcal{V}|}$
end function

RIF approximation (Graph vs MST)





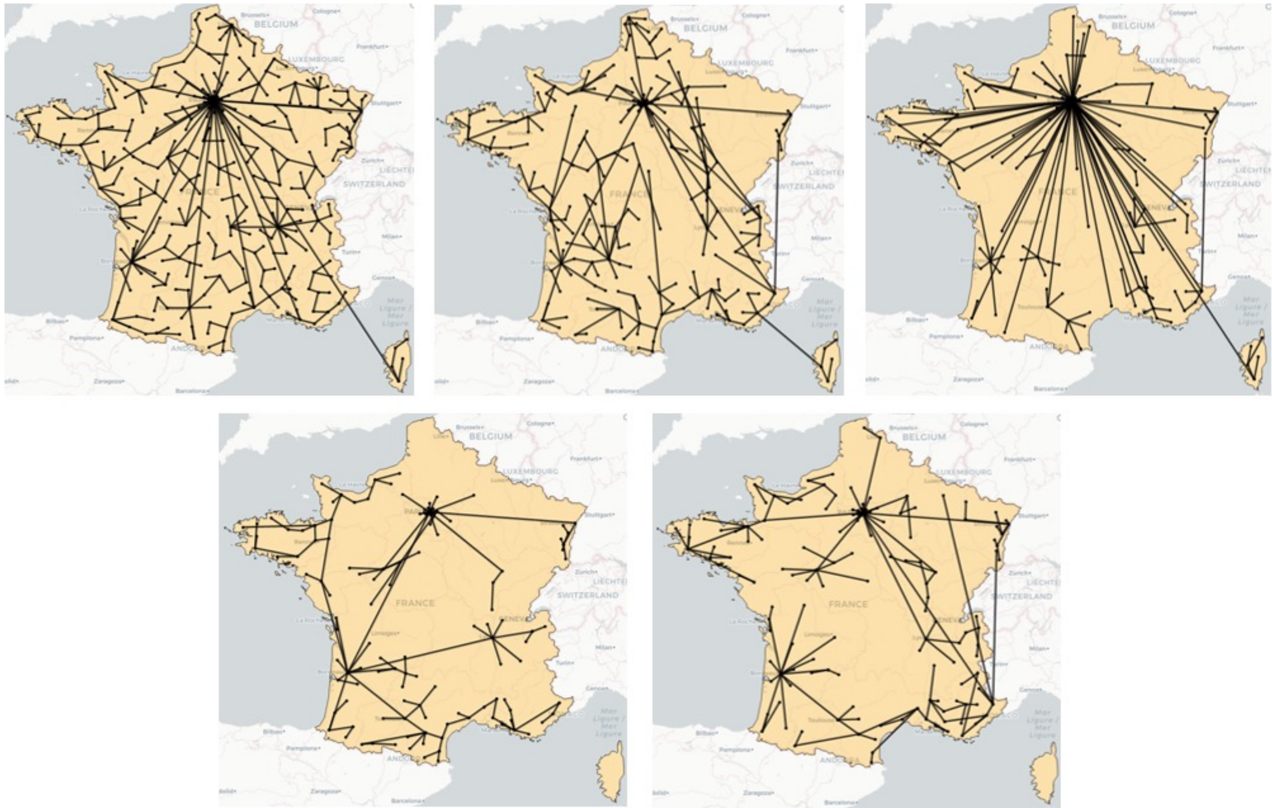
Graphs size	**MSE**	**MAE**	**MAPE**
200 Nodes	6.96×10^{-8}	2.39×10^{-4}	5.53%
500 Nodes	7.20×10^{-9}	7.92×10^{-5}	4.50%
1000 Nodes	5.89×10^{-10}	8.20×10^{-6}	3.92%
2000 Nodes	2.57×10^{-11}	3.32×10^{-6}	3.17%

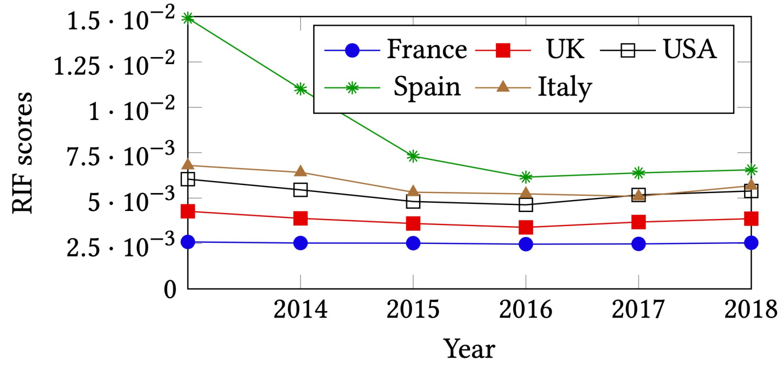
Aggregated circulation graphs

AC **Graphs**	**Aggregation**	**#nodes**	**#edges**
France	Towns	16,824	5,596,001
	Cities	3,678	4,725,402
	Districts	329	801,924
Nouvelle-Aquitaine	Cities	2,665	935,294
	Districts	42	170,403

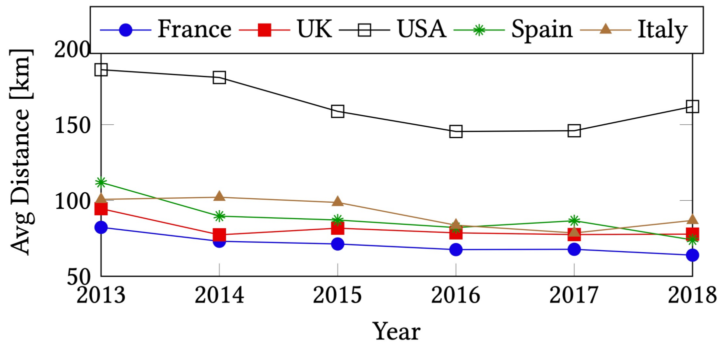
- From 2013 to 2018
- Five different nationalities
- Five levels of aggregation

French, English, American, Spanish and Italian MST in 2018 over France



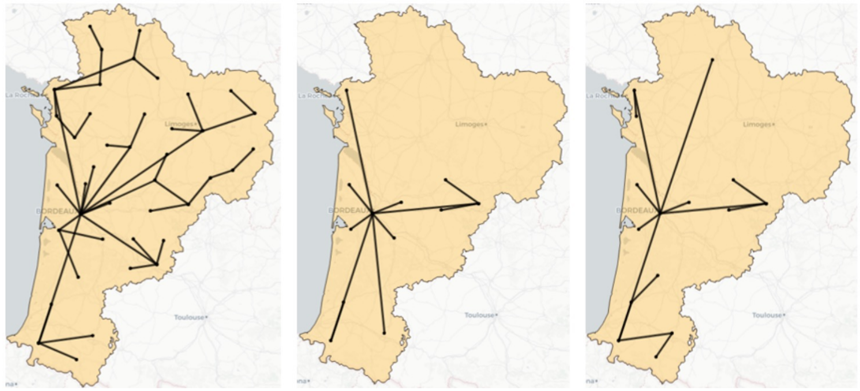


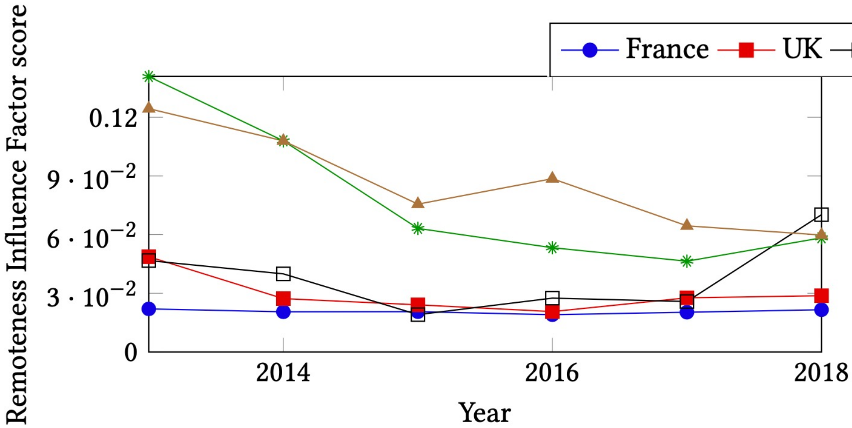
RIF Evolution - France (Town-scale)



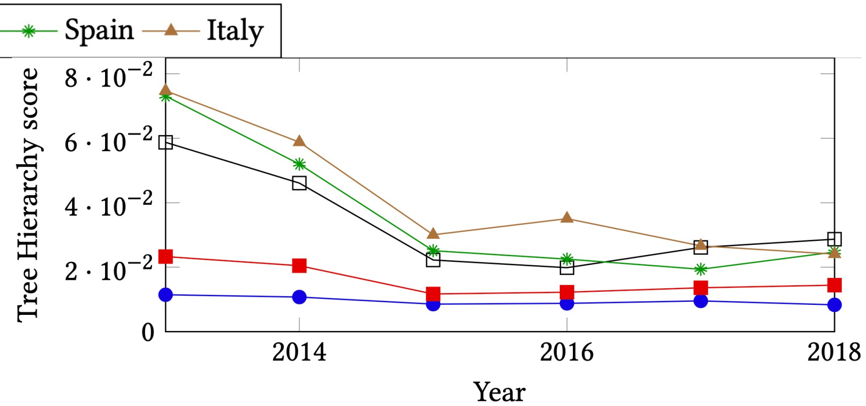
Average geodesic distance (Town-scale)

Experiments - RIF vs Tree Hierarchy





RIF Evolution



Tree Hierarchy evolution

Nouvelle-Aquitaine (District-scale)

Conclusions & perspectives

Propagation Measure on Circulation Graphs for Tourism Behavior Analysis

- An automatic **Maximum Spanning Tree** extraction methodology dedicated to spatiotemporal graphs
- The **Remoteness Influence Factor** (RIF), a new propagation measure
  - Community, year, scale
- Proposed an **optimization strategy** of this computation using Maximum Spanning Trees

Perspectives

- Apply the Remoteness Influence Factor to other fields such as GIS, Traffic Networks or Social Networks
- Study the strength of propagation of tourists in a given area using the principle of percolation
- Study of forests to observe how different tourist regions behave and how they interact within a territory