

Sensor location by joint entropy maximization

Rui S. Shibasaki ^{1,2}, Olivier Péton ^{1,2}, François Queyroi ^{2,3}, Maria I. Restrepo ^{1,2}

¹IMT Atlantique, Nantes, France

²LS2N (UMR CNRS 6004), Nantes, France

³CNRS

EURO, Espoo, July 2022



Table of Contents

- 1 Context
- 2 Sensor Location Problem
- 3 Case Study: Nantes

Context - FLOTES Project

Main goal - Decision Support Tool (DST)

Macro estimation of the freight transportation flow after a modification of the street network

- Note: only specific types of vehicles are accounted for flow measurement
- Users: public authorities

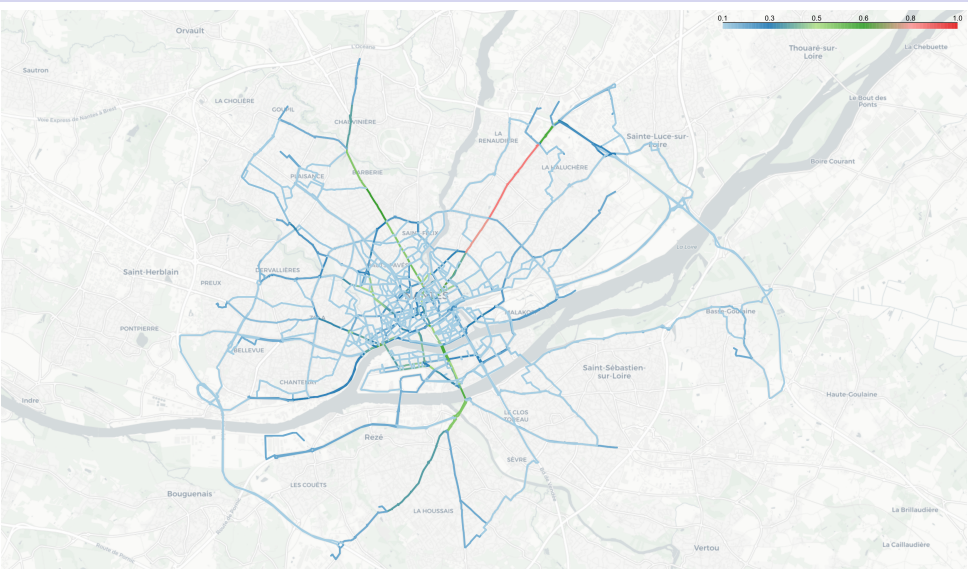
Cities are subject to frequent changes

- New pedestrian areas,
- Blocked areas, limited traffic zones
- New logistic facilities

General approach

- Mathematical models for estimations
- Verify/correct estimations using real data collected by sensors

Freight transportation flow estimation



Freight transportation flow estimation

Input Data

- Graph (city's street network)
- Client and depot points
- Clients' demands and their assignments (depot)
- Demand quantities are uncertain.

Algorithm 1: Main algorithm

- 1 Define demand scenarios
 - 2 **while** *stopping criteria not reached* **do**
 - 3 Simulate the flow in the different scenarios (Traffic Map)
 - 4 Determine best-fit scenarios according to real data
 - 5 Discard unreal scenarios and derive others
 - 6 Network modification: apply changes to the graph.
 - 7 Define an average traffic map considering remaining scenarios
-

Scenarios & Simulation

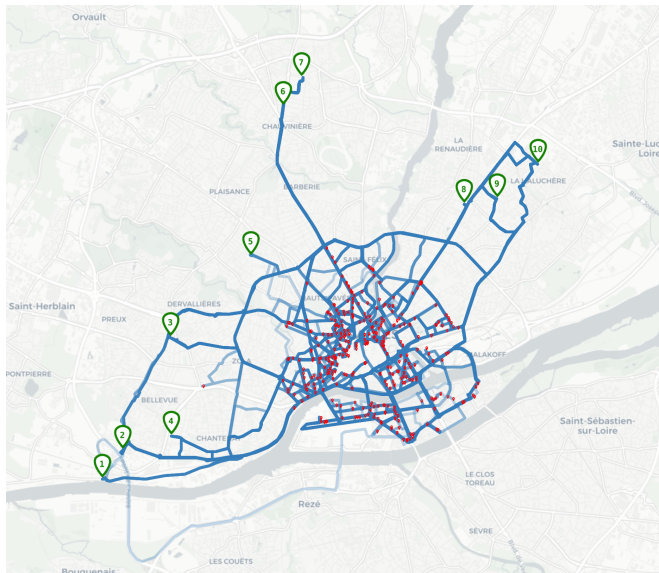
Scenarios

- Demands are grouped by type
- Quantity of each demand type i in scenario k : $X_{ki} \sim N(\mu_{ki}, \sigma_{ki})$

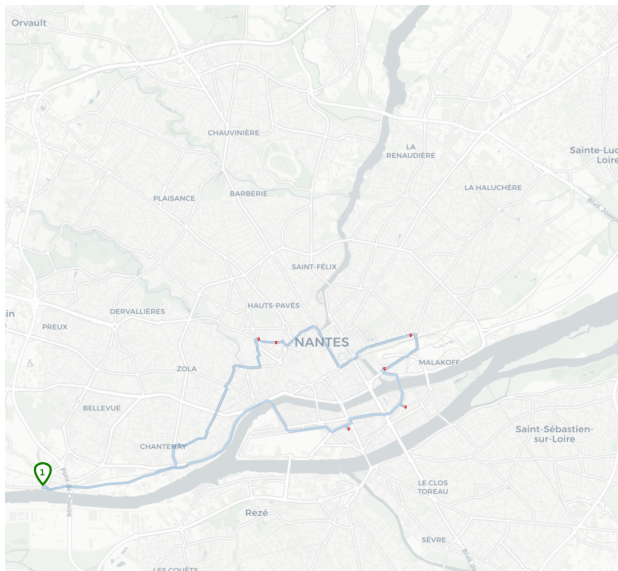
Freight transportation flow simulation

- Several Vehicle Routing Problem (VRP) instances.
- Demand quantities randomly generated according to the scenarios
- Solved by Large Neighborhood Search Metaheuristic.
- Flow: number of tours routed on an arc.

Example of simulated flow (Traffic Map)



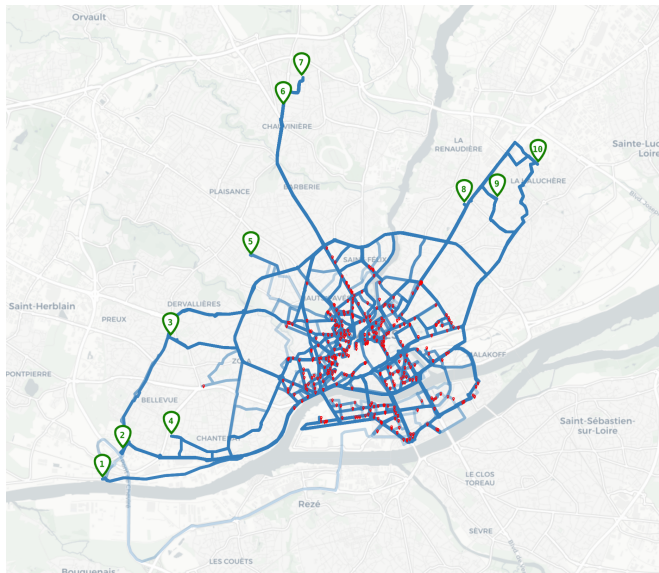
Example of simulated flow (Traffic Map)



Example of simulated flow (Traffic Map)



Example of simulated flow (Traffic Map)



Best-fit scenarios according to real data

- From simulations, we have flow data for each scenario. Which ones would best represent reality?

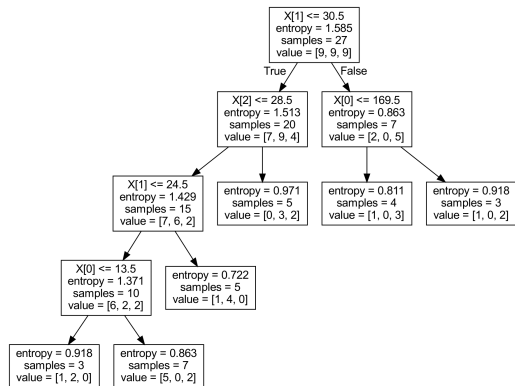
Sensor locations

- Place sensors.
- Use the simulation data on the sensor locations to train a Machine Learning algorithm (Decision Trees).
- Use the real data captured by the sensors in their locations to predict (in probabilities) the scenario of the real flow.

Decision Trees

- s: simulation, k: scenario
- 9 simulations \times 3 scenarios
- 27 lines

node 0	node 1	node 2	scenario
$F_{sk}(0)$	$F_{sk}(1)$	$F_{sk}(2)$	
10	0	5	1
20	5	10	1
15	13	20	1
20	25	23	2
30	25	26	2
25	30	31	2
100	25	50	3
200	50	41	3
150	45	47	3
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮



Define a schematic way of reasoning for prediction, based on training data.

Decision Trees - Predict probabilities

- Flow measured
 by sensors:

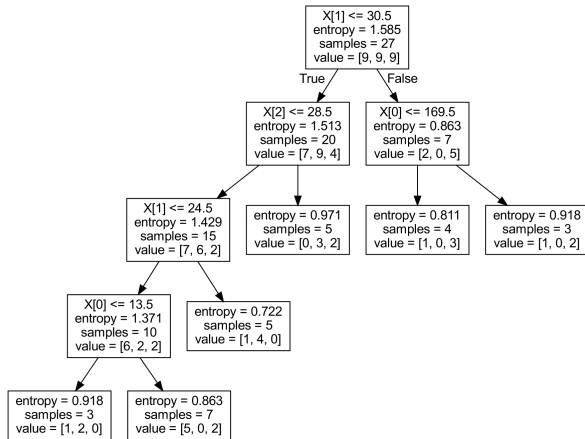
node 0	node 1	node 2
$\bar{F}_{sk}(0)$	$\bar{F}_{sk}(1)$	$\bar{F}_{sk}(2)$
56	10	50

- Prediction:

scenario 1: 0% (0/5)

scenario 2: 60% (3/5)

scenario 3: 40% (2/5)



Sensor location

Two approaches

Sensor Location by Maximum Entropy Sampling (future research)
Sensor Location by Feature Selection

- Feature Selection **X** Maximum Entropy Sampling

Maximum Entropy Sampling

- The flow data on selected nodes tend to provide equal probabilities among scenarios
- A second problem: given the real flow, how to build a demand scenario to be tested?

Feature Selection

- The flow data on selected nodes tend to better distinguish scenarios

Sensor Location by Feature Selection

- Provide the entire simulation data to train the Decision Tree.
- The selection of nodes will be naturally done by feature selection.
- Place sensors on nodes used in branching.

node 0	...	node 100	...	scenario
$F_{sk}(0)$...	$F_{sk}(100)$...	
10	...	5	...	1
20	...	10	...	1
15	...	20	...	1
20	...	23	...	2
30	...	26	...	2
25	...	31	...	2
100	...	50	...	3
200	...	41	...	3
150	...	47	...	3
⋮	⋮	⋮	⋮	⋮

Limited number of sensors - Regularized Decision Trees

- The Decision Tree cannot select more than c features (nodes)

Case Study: Nantes

Graph

- OpenStreetMap
- OSMnx ¹

Depot/clients locations

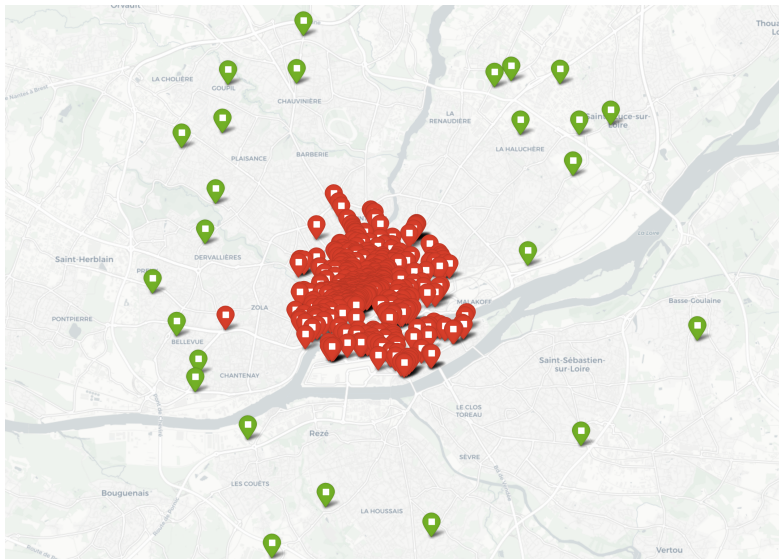
- List of adresses OpenData: <https://data.nantesmetropole.fr/>
- Relate it to the graph nodes

Demands

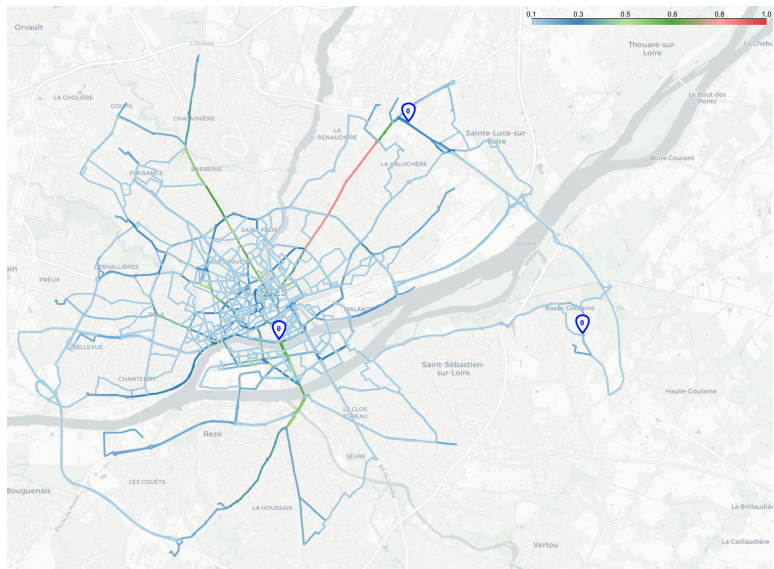
- Random client-depot assignment.
- The types of demands of each client were randomly chosen
- The depot serving the demand was also randomly chosen

¹Boeing, G. 2017. "OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks." *Computers, Environment and Urban Systems*. 65, 126-139.

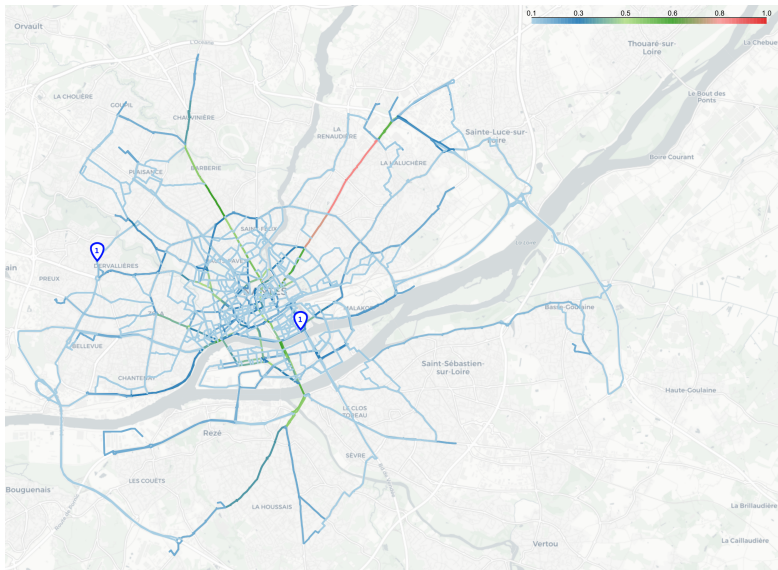
Locations of 444 clients and 25 depots



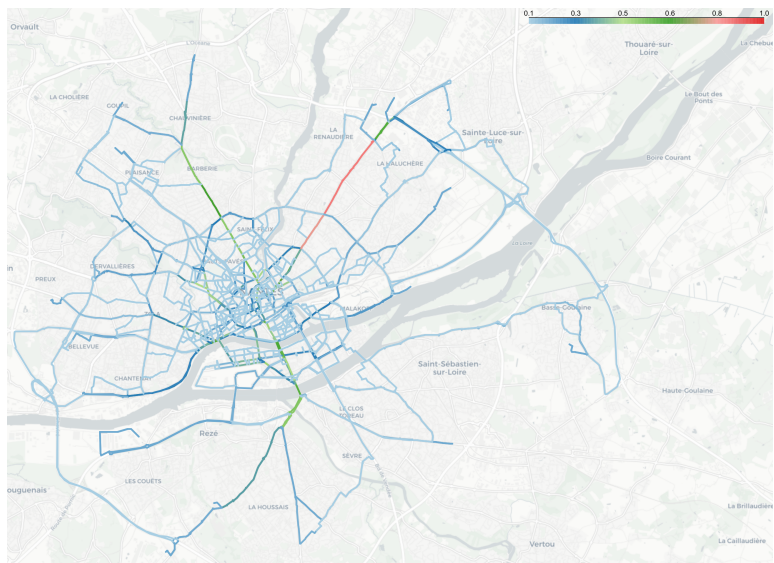
First iteration



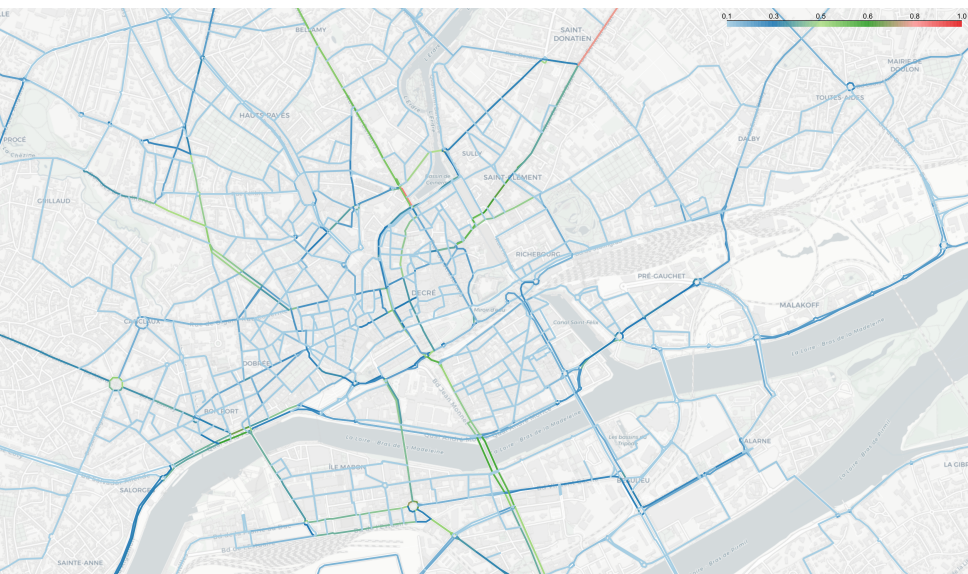
Second iteration



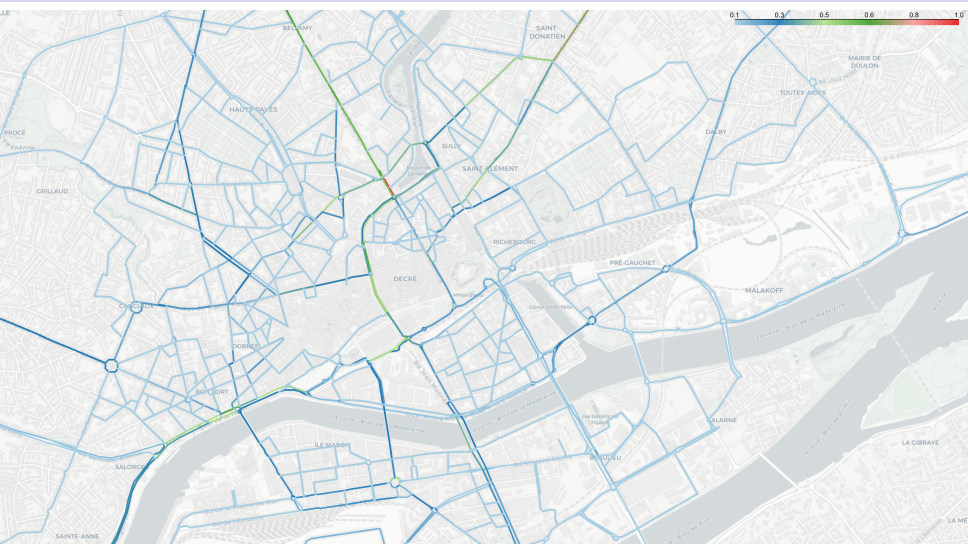
Final flow



Pedestrian Zone



Pedestrian Zone



Computational experiments

4 types of demands

10 scenarios x 10 simulations per scenario

5 runs per instance

Instance	Real	Avg Dev	Max Dev	Max JSDiv	Iters	Sensors
CF_1	No	0.37	3.3	0.0	2	6
CF_2	No	0.37	3.5	0.1	2	6
CF_3	No	0.38	3.5	0.0	3	6
CF_1	Yes	0.32	2.9	0.0	3	7
CF_2	Yes	0.36	3.0	0.0	2	6
CF_3	Yes	0.34	3.0	0.0	3	7
F_1	No	2.16	18.6	175.8	1	3
F_2	No	1.32	10.5	95.0	1	3
F_3	No	2.83	22.9	63.3	2	4
F_1	Yes	0.32	2.9	0.0	2	5
F_2	Yes	0.34	3.0	0.0	1	3
F_3	Yes	0.31	2.7	0.0	1	3

Dev: $|F(a) - \bar{F}(a)| / \max_{a \in A} \{\bar{F}(a)\}$

JSDiv : Jensen-Shannon Divergence

Computational experiments

12 types of demands

10 scenarios x 10 simulations per scenario

5 runs per instance

Instance	Real	Avg Dev	Max Dev	Max JSDiv	Iters	Sensors
CF_1	No	0.54	5.9	0.1	4	10
CF_2	No	0.53	4.9	0.1	4	11
CF_3	No	0.55	4.7	0.1	3	9
CF_1	Yes	0.52	5.5	0.0	4	12
CF_2	Yes	0.54	5.2	0.1	3	8
CF_3	Yes	0.53	4.7	0.1	5	13
F_1	No	0.80	6.2	50.0	2	5
F_2	No	1.65	12.2	50.0	2	5
F_3	No	2.98	21.5	105.4	3	9
F_1	Yes	0.59	5.5	4.1	3	8
F_2	Yes	0.53	4.4	0.0	1	3
F_3	Yes	0.51	4.7	0.0	1	3

Dev: $|F(a) - \bar{F}(a)| / \max_{a \in A} \{\bar{F}(a)\}$

JSDiv : Jensen-Shannon Divergence

Conclusion

Macro information about freight transportation flow

- Where the vehicles pass the most
- How the flow reacts to modifications in the network

Some assumptions have been made

- An idea of what is the real demand scenario is considered to exist
- Locations and assignments of clients/depots are well determined

Future research

- How to determine scenarios given some real flow data
- How to generalize the model: uncertain assignments client-depot