Lab in Data Science

Prof. Verscheure

SBB Route Planner 2.0

Alexandre Poussard, Olivier Couque, Romain Choukroun

June 13, 2018

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Outline:

- Motivation
- Network Modelisation
  - Assumptions
  - Method
  - Final graph
- Routing Algorithm
  - Needs
  - Breadth-First-Search
  - Dijkstra
- Delay-based Model
- Visualization
- Pros & Cons

<u>Motivation:</u>

- Today's apps: <u>**theoretical**</u> quickest path
- What's new?

Big Data!

- Goal: flexibility & simplicity to customers

June 13, 2018

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Network Modelisation:

Assumptions:

- Change of transportation within a station takes 2 mins
- Maximum distance to walk is 500m
- Walking speed of 60m/min
- 2 types of schedule: working days & weekend

Network Modelisation:

Process:

- Remove rare / unfinished stops
- Filter stations in 10km range & match them with SBB
- Defined realistic schedule: weekdays + weekend
- Modelisation of trips (all stations along one path)

Network Modelisation:

Baseline Graph:

- Big dataframe containing every edges of every day

- So what next ?

	trip_id	departure	departure_time	arrival	arrival_time	time
0	85:11:13752:001	Zürich Stadelhofen	1900-01-01 01:52:00	Zürich HB	1900-01-01 01:55:00	3.0
1	85:11:13752:001	Zürich HB	1900-01-01 01:57:00	Zürich Hardbrücke	1900-01-01 01:59:00	2.0
2	85:11:13752:001	Zürich Hardbrücke	1900-01-01 01:59:00	Zürich Altstetten	1900-01-01 02:01:00	2.0
3	85:11:13752:001	Zürich Altstetten	1900-01-01 02:01:00	Urdorf	1900-01-01 02:06:00	5.0
4	85:11:13752:001	Urdorf	1900-01-01 02:06:00	Urdorf Weihermatt	1900-01-01 02:07:00	1.0

Routing Algorithm:

Needs:

- Freeze the situation
- Graph !
- No delay yet

Routing Algorithm:

Freezing the graph:

- Give us a context: time, place...
- Breadth-First Search
- Walking times !

Routing Algorithm:

Routing:

- Multitude of routing algorithms
- Focus on the effect of the delays
- Run Dijkstra !

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Delay-Based Model:

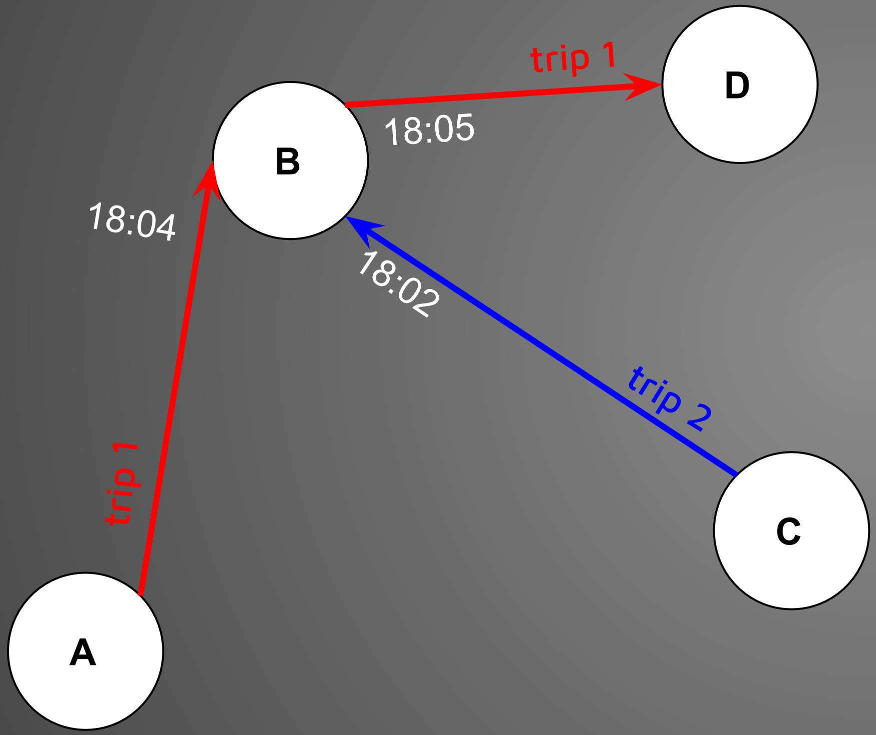
- Exponential Distribution: $$F(x;\lambda) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0, \\ 0 & x < 0 \end{cases}$$

Cumulative function

- **λ = median delay** -> trip 2: [0, 1, 1, 2, 3, 3, 3, 5, 8] -> λ = ⅓

- Compute the success probability of a connection at each station

Delay-Based Model:



Probability of connection success at B

trip 2: p = F(delay=3; λ=1/3) = 0.64
trip 1: p = 1 because same trip

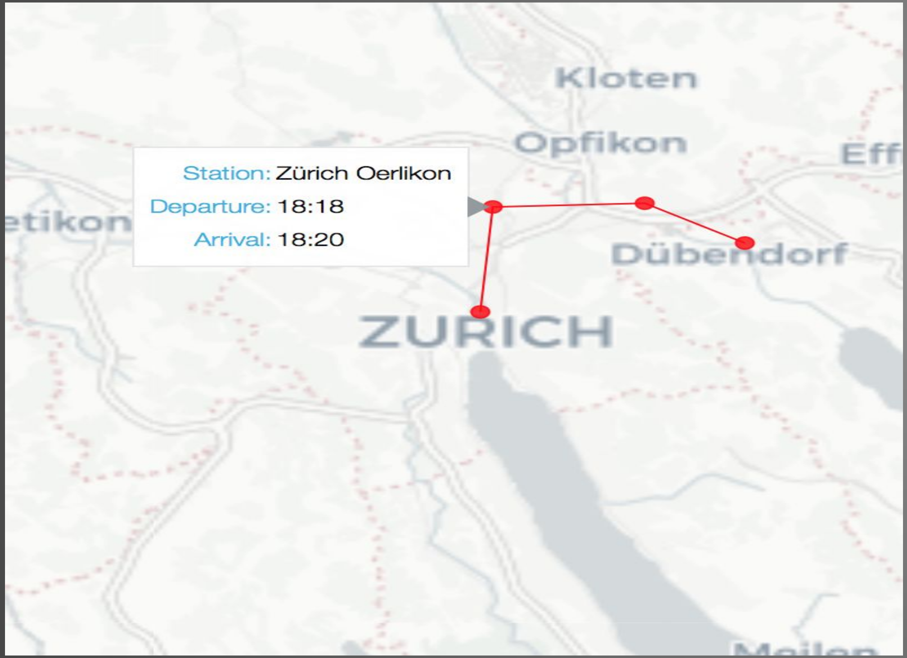
Station B

(trip 1, trip 1) : p = 1.0
(trip 2, trip 1) : p = 0.64
...

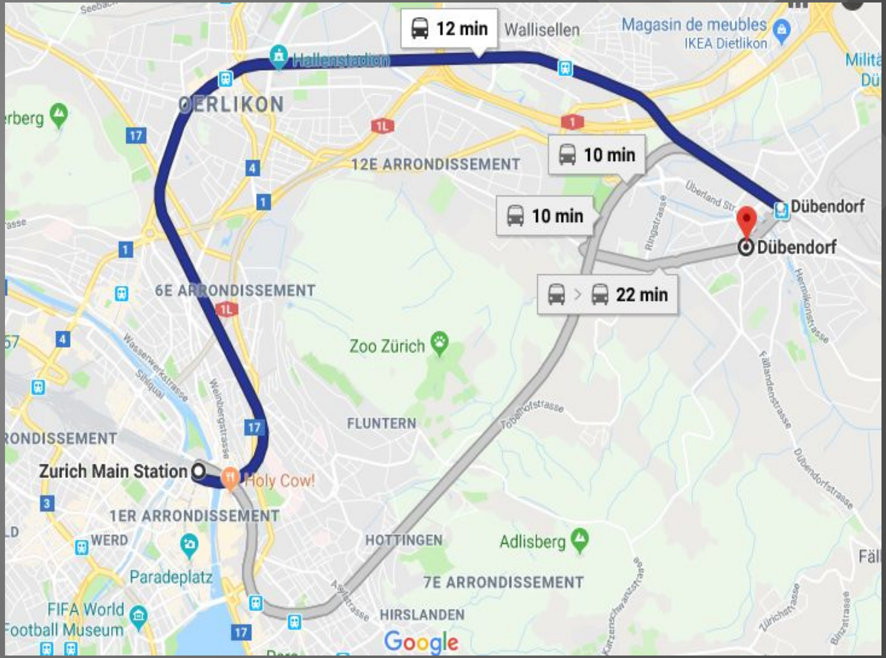
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Visualization: Shortest path comparison

Our model: Zürich HB - Dübendorf: **20 minutes**



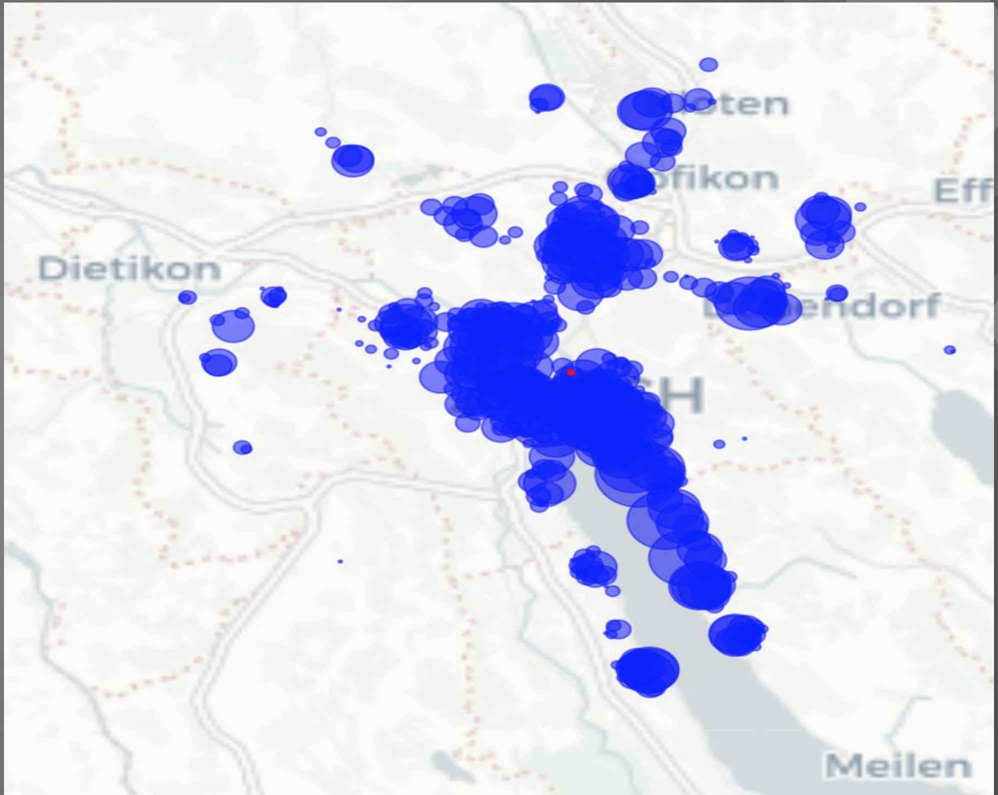
Google Maps: Zürich HB - Dübendorf: **12 minutes**



June 13, 2018

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Visualization: Isochronous Map



User query

From: Zürich HB

Departure Time: 18h00

Travel duration: 30 minutes

Pros & Cons:

Pros

- We consider the possible delays a trip can have
- Better consideration of walking times between stations
- Could offer the possibility for the user and the company to see which trips tend to be late

Cons / Improvements

- Takes some time to run a user's query (around 1 minute)
- Users aren't able to defer their decisions and adapt their journey
- Didn't take into account the rush hour factor

Thank You !

Questions ?

June 13, 2018

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE