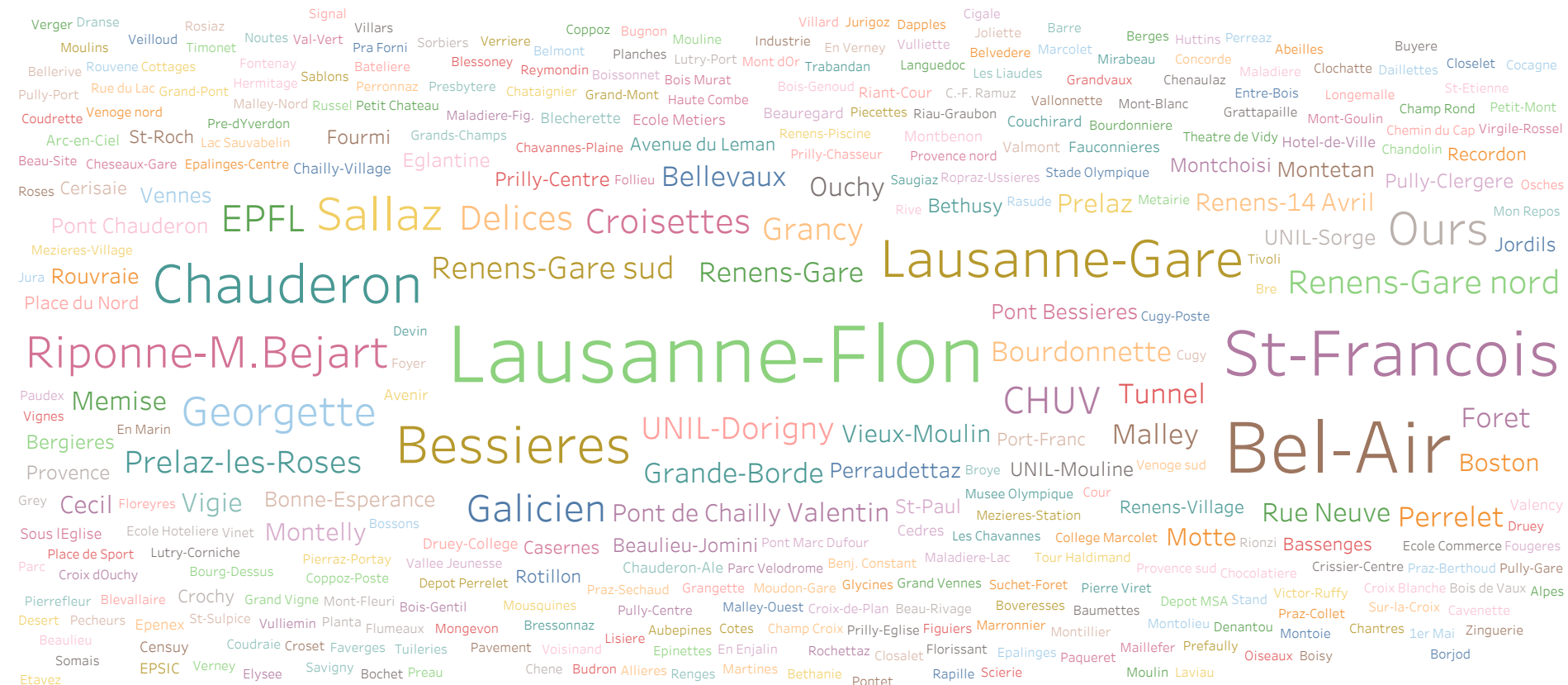




Getting data about controller movement

Mine data from Twitter feed of @BustedLAUSANNE and Busted-App

```
INFO: Starting new HTTPS connection (1): twitter.com
INFO: Got 12658 tweets for BustedLausanne%20until%3A2017-06-24%20until%3A2017-06-23.
INFO: Got tweets ranging from 2013-10-22 to 2017-06-22
INFO: Running iteration no 4, query is 'BustedLausanne until:2017-06-24 until:2017-06-23 until:2013-10-23'
INFO: Querying BustedLausanne until:2017-06-24 until:2017-06-23 until:2013-10-23
INFO: Starting new HTTPS connection (1): twitter.com
INFO: Got 13517 tweets (1 new).
INFO: Starting new HTTPS connection (1): twitter.com
INFO: Got 1 tweets for BustedLausanne%20until%3A2017-06-24%20until%3A2017-06-23%20until%3A2013-10-23.
INFO: Got tweets ranging from 2013-10-22 to 2013-10-22
INFO: Running iteration no 5, query is 'BustedLausanne until:2017-06-24 until:2017-06-23 until:2013-10-23 until:2013-10-22'
INFO: Querying BustedLausanne until:2017-06-24 until:2017-06-23 until:2013-10-23 until:2013-10-22
INFO: Starting new HTTPS connection (1): twitter.com
INFO: Got 0 tweets for BustedLausanne%20until%3A2017-06-24%20until%3A2017-06-23%20until%3A2013-10-23%20until%3A2013-10-22.
hec-170-69:Desktop Vaibhav$
```

Data Format



Busted Lausanne @BustedLAUSANNE · 20h



#Controleurs Lausanne-Flon, ligne m2 → Croisettes (via busted-app.com)

 Translate from French



Busted Lausanne @BustedLAUSANNE · 5h



#Controleurs Bel-Air, hors du véhicule (via busted-app.com)

 Translate from French

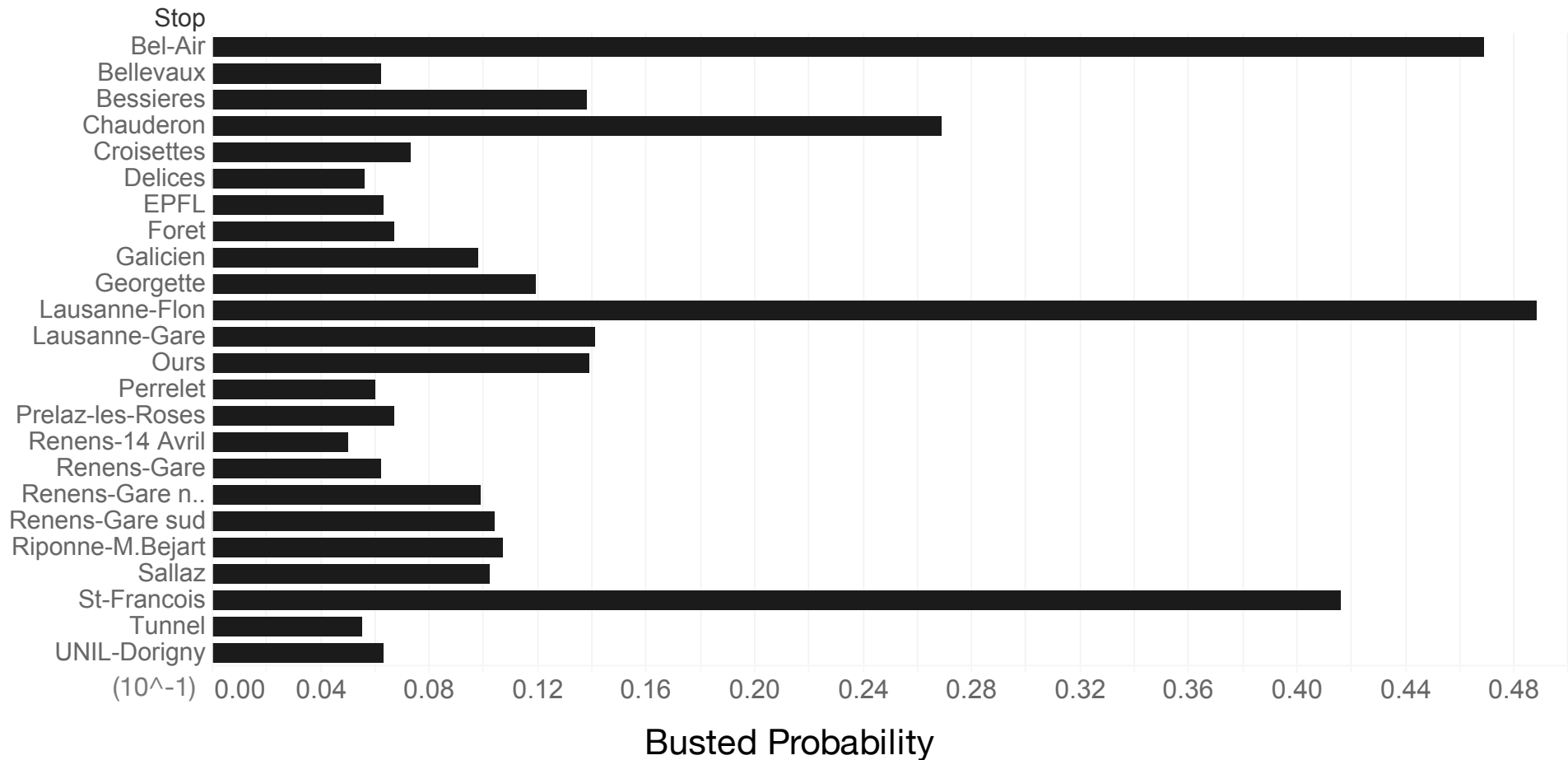


Format: [Stop], [line to destination]/[outside/inside the vehicle], [timestamp]

Analyze the data to check for attack surfaces

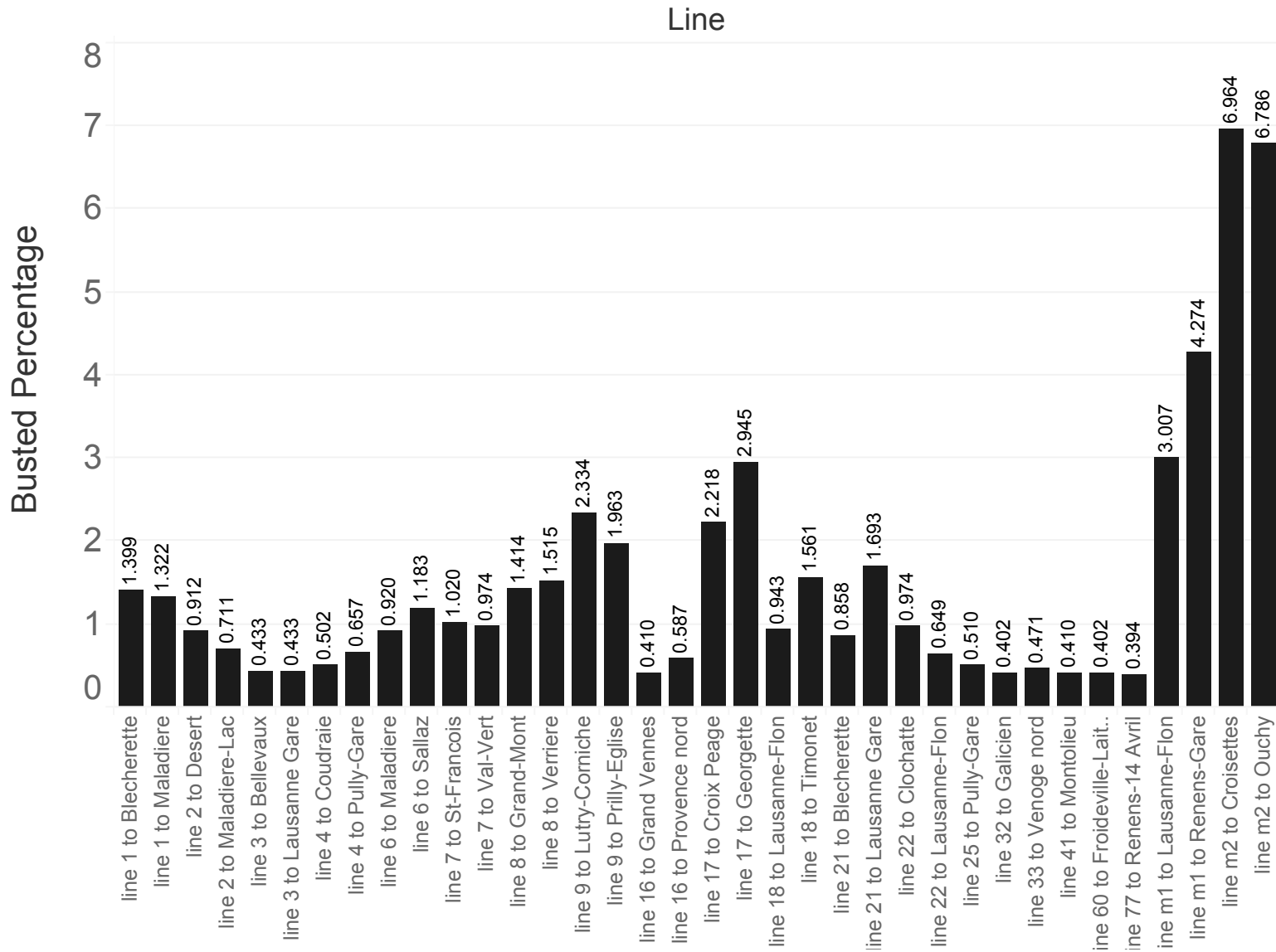
- What stops/lines come under high-frequency control?
- At what hours of the day is controlling prominent?
- Is there any correlation between the stop/line and the control time?
- At which stops is controlling performed outside the vehicle?
- Given the current controller position, is the next position predictable?
- Any observable patterns in weekday/weekend/seasonal regularities?

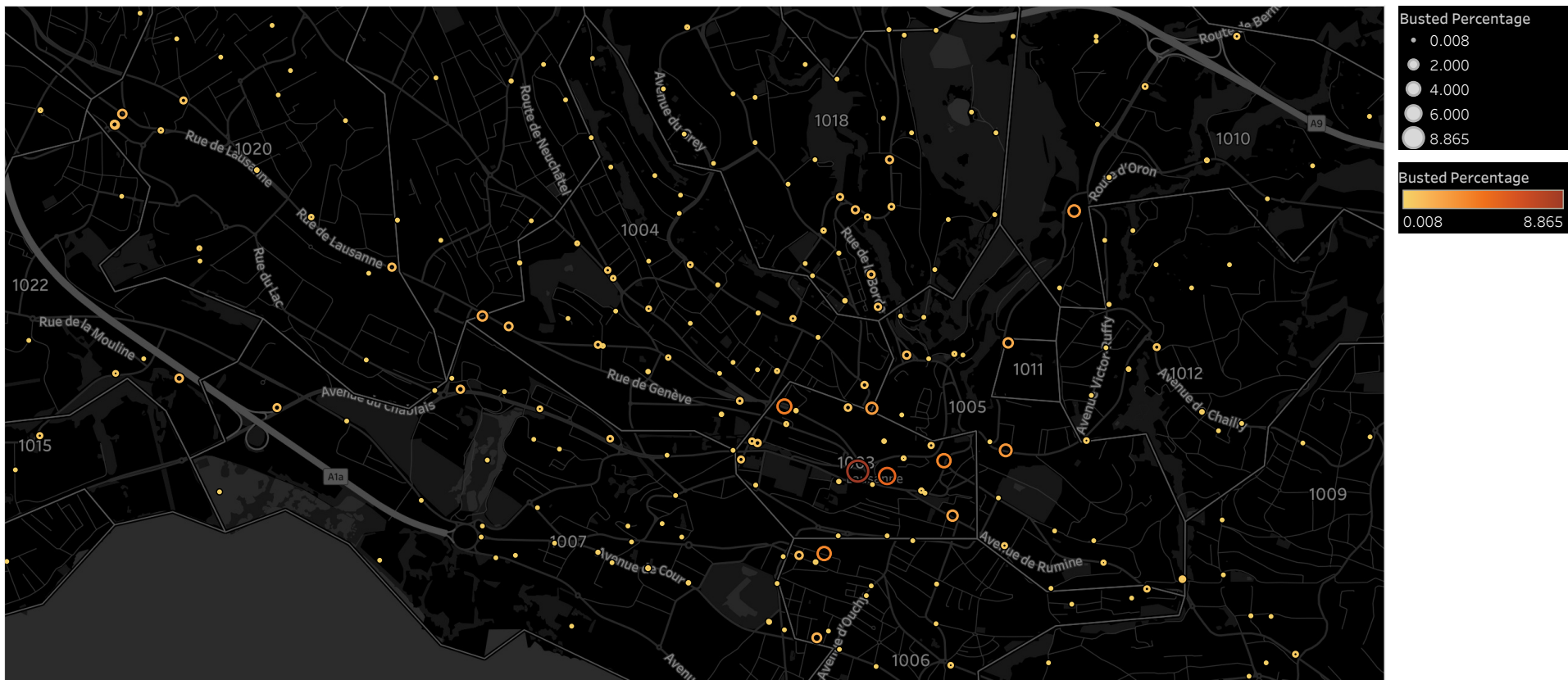
Stop against “busted” probability



*busted probability at all other stops < 0.005

Line v/s. busted probability



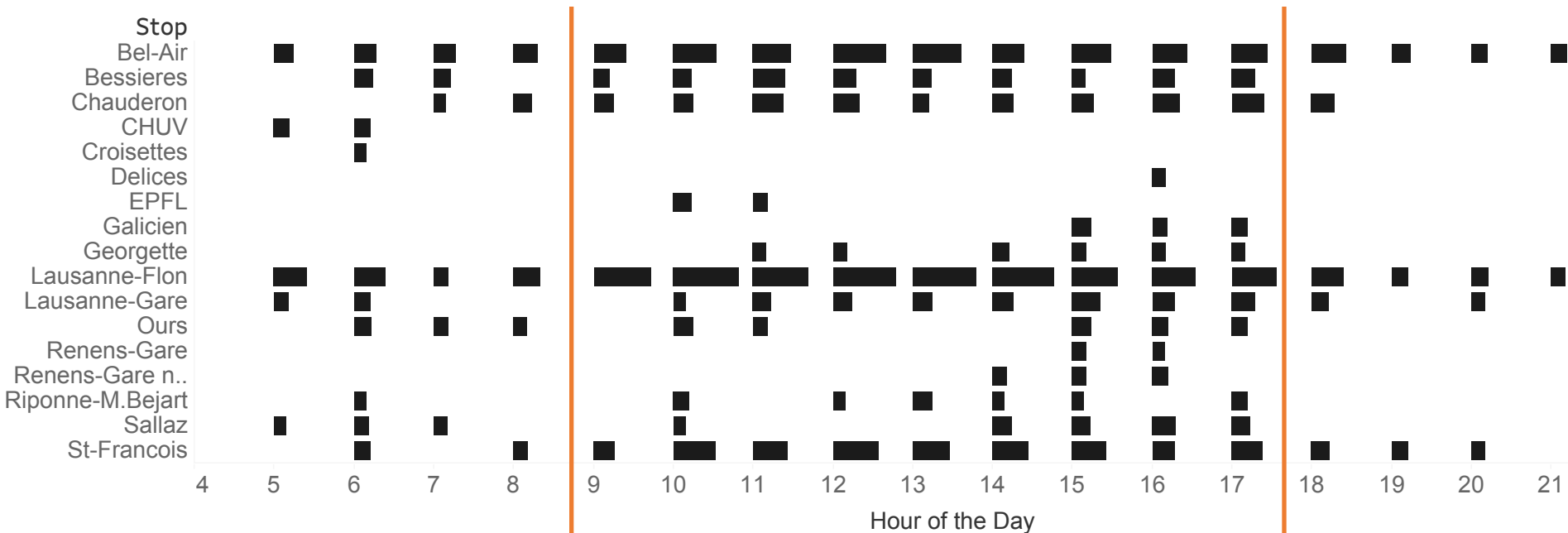


Zip-code area v/s. busted percentage

Attack surface

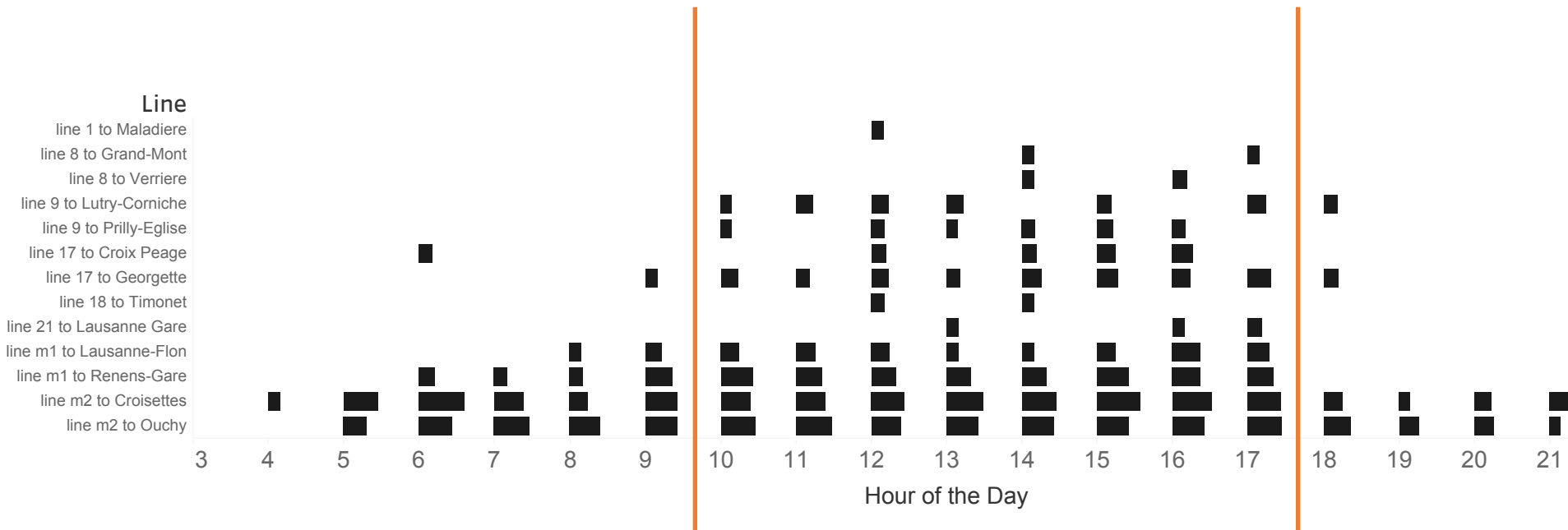
- If I know that a stop/line is controlled at a high frequency I will always buy a ticket while going there.
- **However, I will not buy a ticket at all other stops.**

Hour of day v/s. busted probability



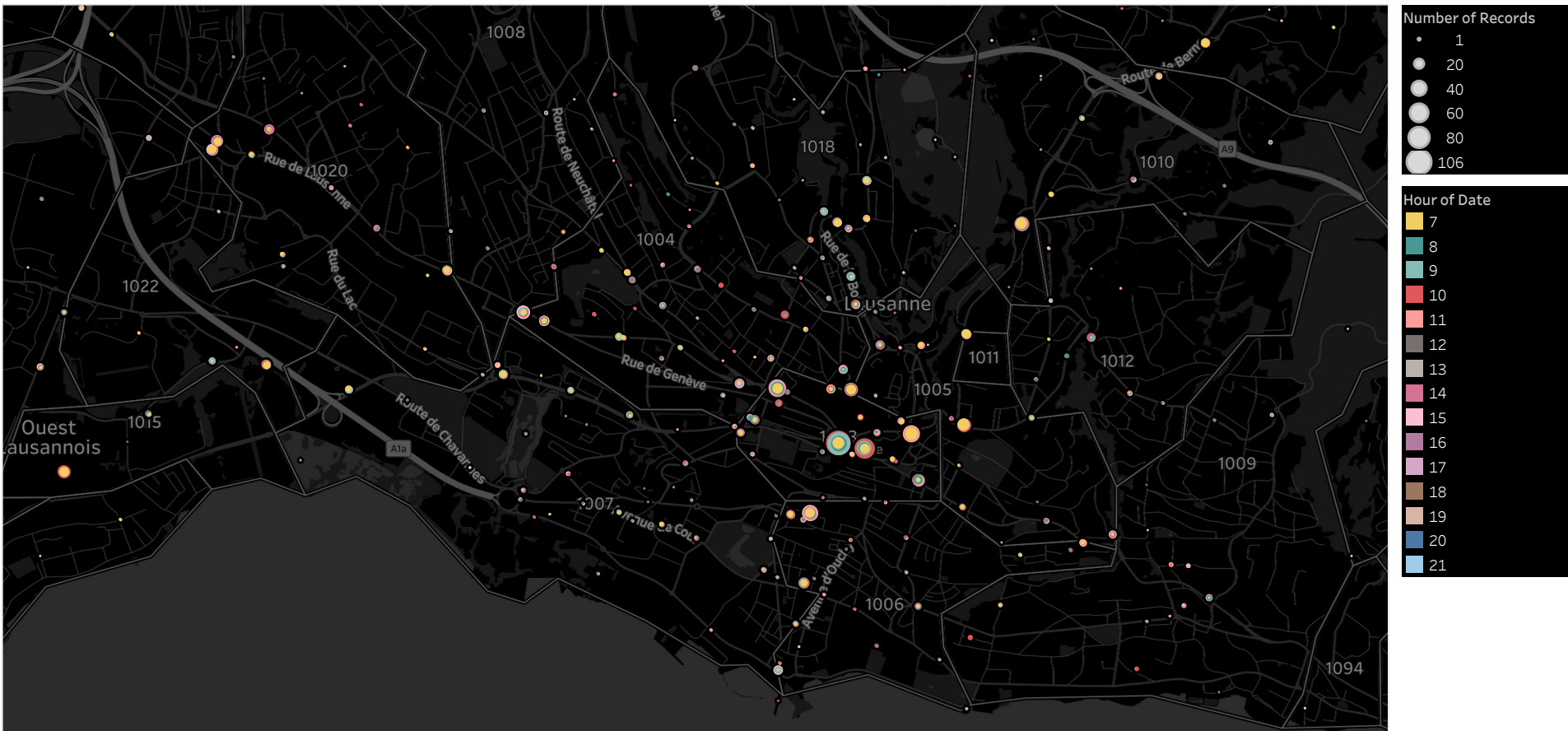
Always buy a ticket at the above stops at these times and vice versa

Line v/s hour of the day v/s busted probability

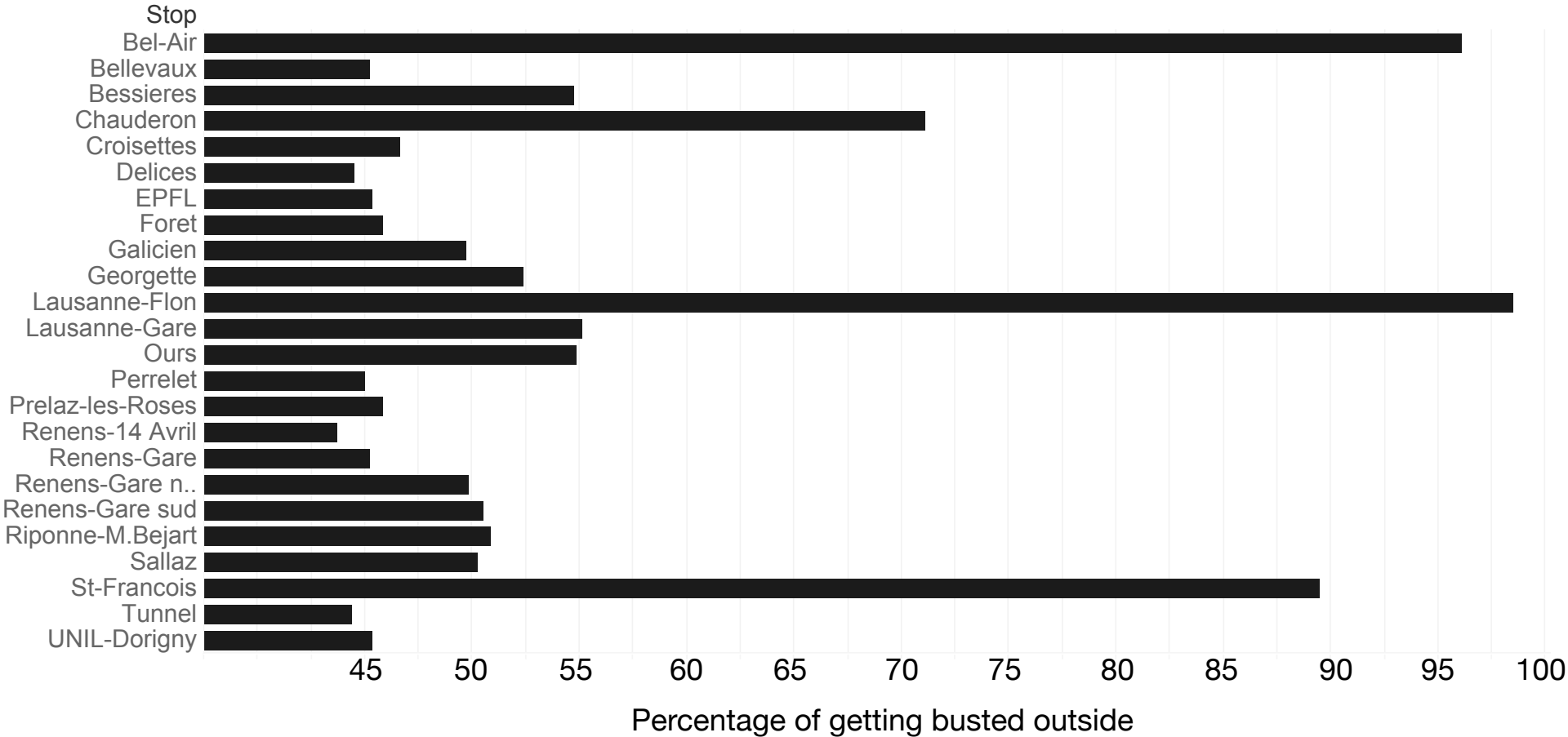


Always buy a ticket at the above lines at these times and vice versa

Area v/s. day time v/s. busted probability



Probability of getting busted outside the vehicle



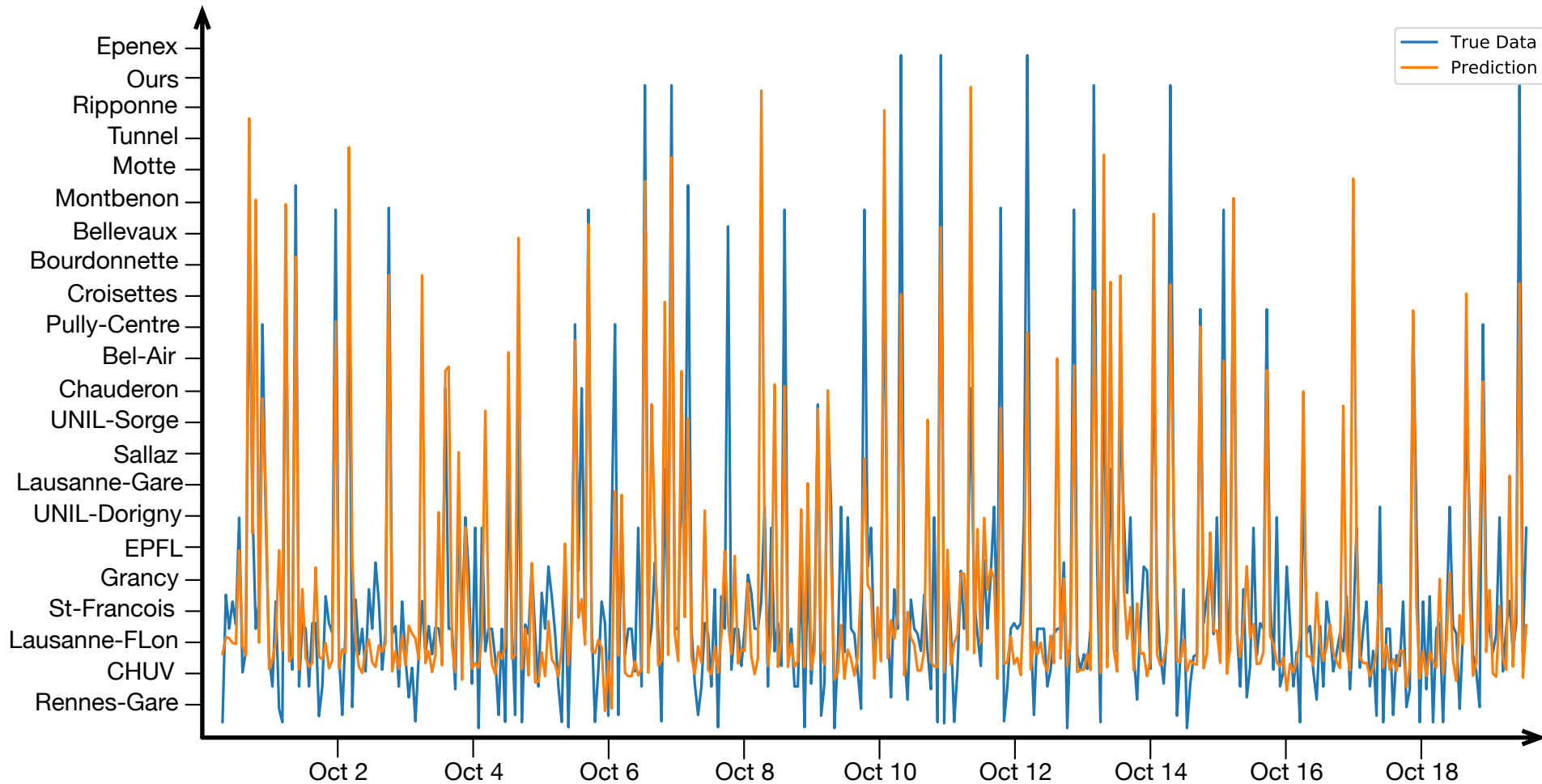
If you don't have a ticket, don't get out at these stops. Wait for the next one

Movement pattern



More than 70% confidence to predict where the controllers will go next at certain stops

Overall Prediction



More than 80% confidence to predict where the controllers will go next

Several repeatable patterns discovered

Conclusion

- Several attack surfaces available to essentially travel for free!
- Can this be avoided ?

Measures and data requirement

- Using a model which is randomized in space and time, adhering to some constraints
 - viz. some stops must always be monitored.
- Data required
 - Passengers per line at all times of the day
 - Number of of subscriptions (monthly, yearly) per zone
 - Number of paper tickets sold per bus stop
 - Number of SMS tickets sold per day and per location
 - Number of busted people per day and per location

Organization of the controllers

- Fixed salaries?
- Commission based salaries?
- Range of salaries?
- Itinerary?