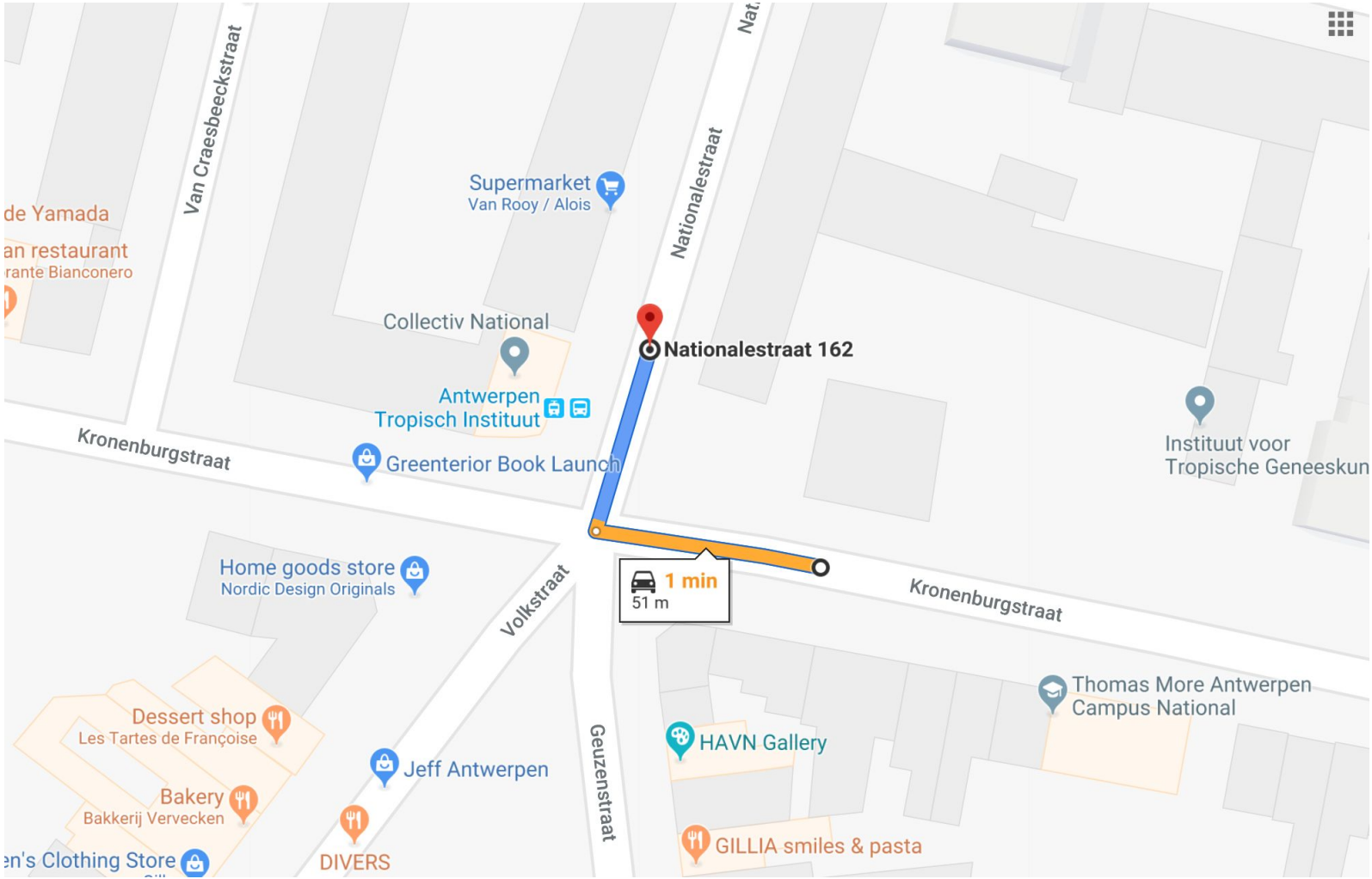
Predicting phase durations of traffic lights using live Open Traffic Lights data

9th September 2019
International Workshop On Semantics For Transport 2019

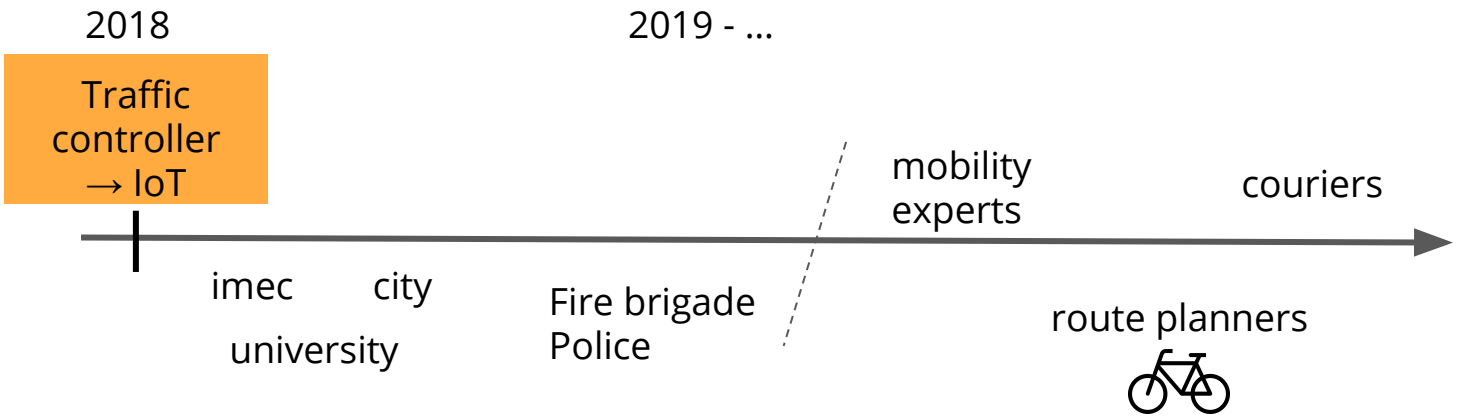
Brecht Van de Vyvere – Karel D'Haene - Kurt D'Haene - Pieter Colpaert – Ruben Verborgh
IDLab (IMEC – Ghent University)

imec

UNIVERSITEIT GENT


Van Craesbeeckstraat
Nationalestraat
Supermarket
Van Rooy / Alois
de Yamada
an restaurant
rante Bianconero
Collectiv National
Nationalestraat 162
Antwerpen
Tropisch Instituut
Kronenburgstraat
Greenterior Book Launch
Instituut voor
Tropische Geneeskun
Home goods store
Nordic Design Originals
1 min
51 m
Kronenburgstraat
Volkstraat
Thomas More Antwerpen
Campus National
Dessert shop
Les Tartes de Françoise
Jeff Antwerpen
HAVN Gallery
Geuzenstraat
Bakery
Bakkerij Vervecken
DIVERS
GILLIA smiles & pasta

Innovating traffic control data



How can route planners reuse this information?

Can the phase duration of dynamic traffic lights be predicted accurately and how should this be done?

Departure: Kronenburgstraat (> nr. 48) ▾ Arrival: Nationalestraat ▾

> 6 seconds

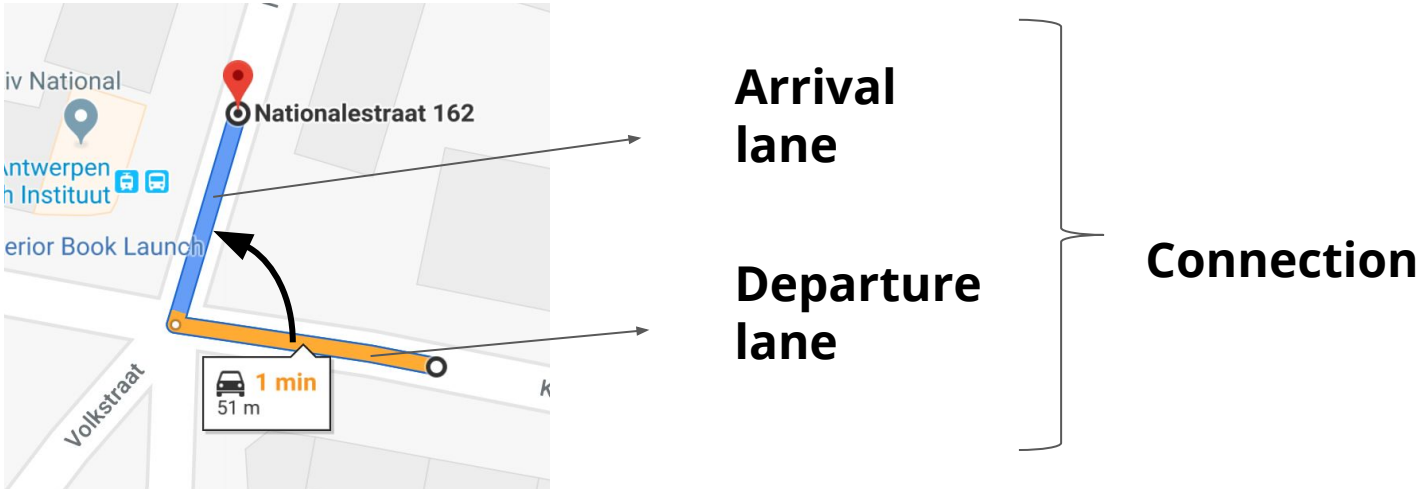
Simple algorithms give a good start, but more context information is needed

Semantics of traffic lights data

Predicting phase durations

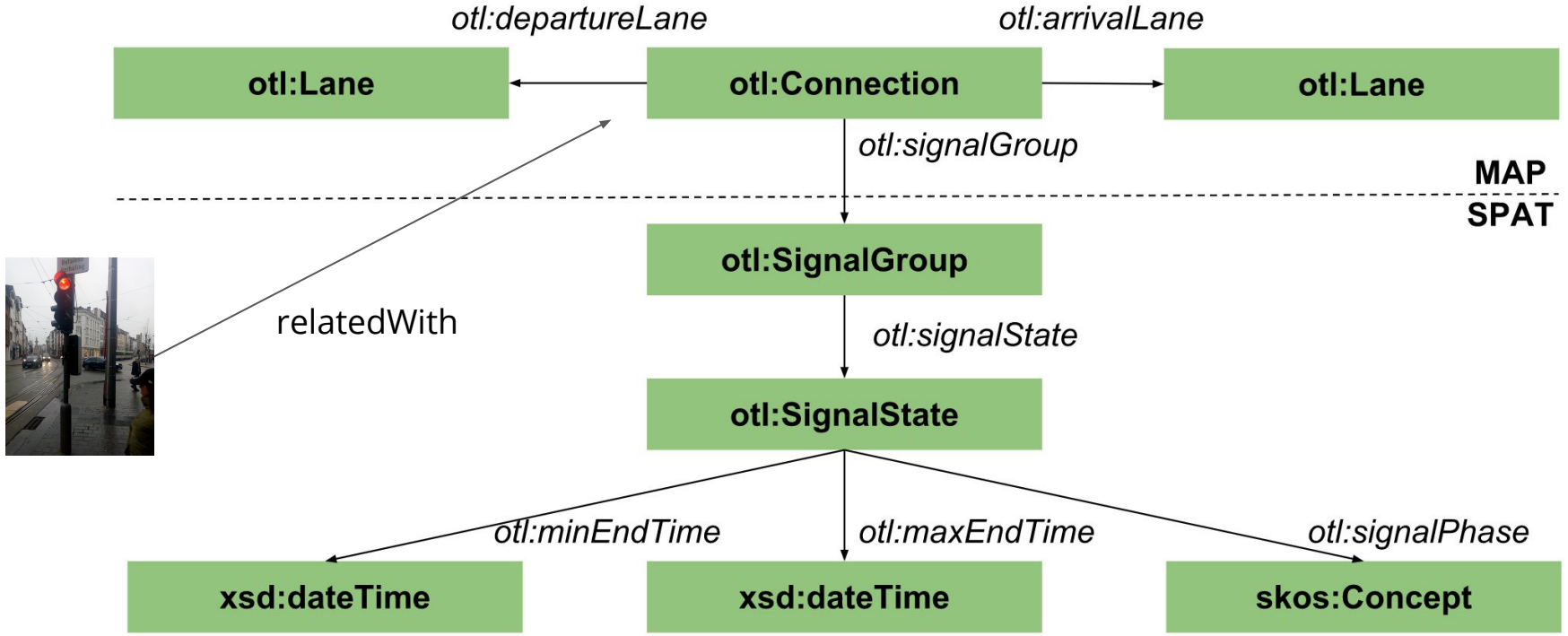
Demo

MAP describes which connections are possible at an intersection

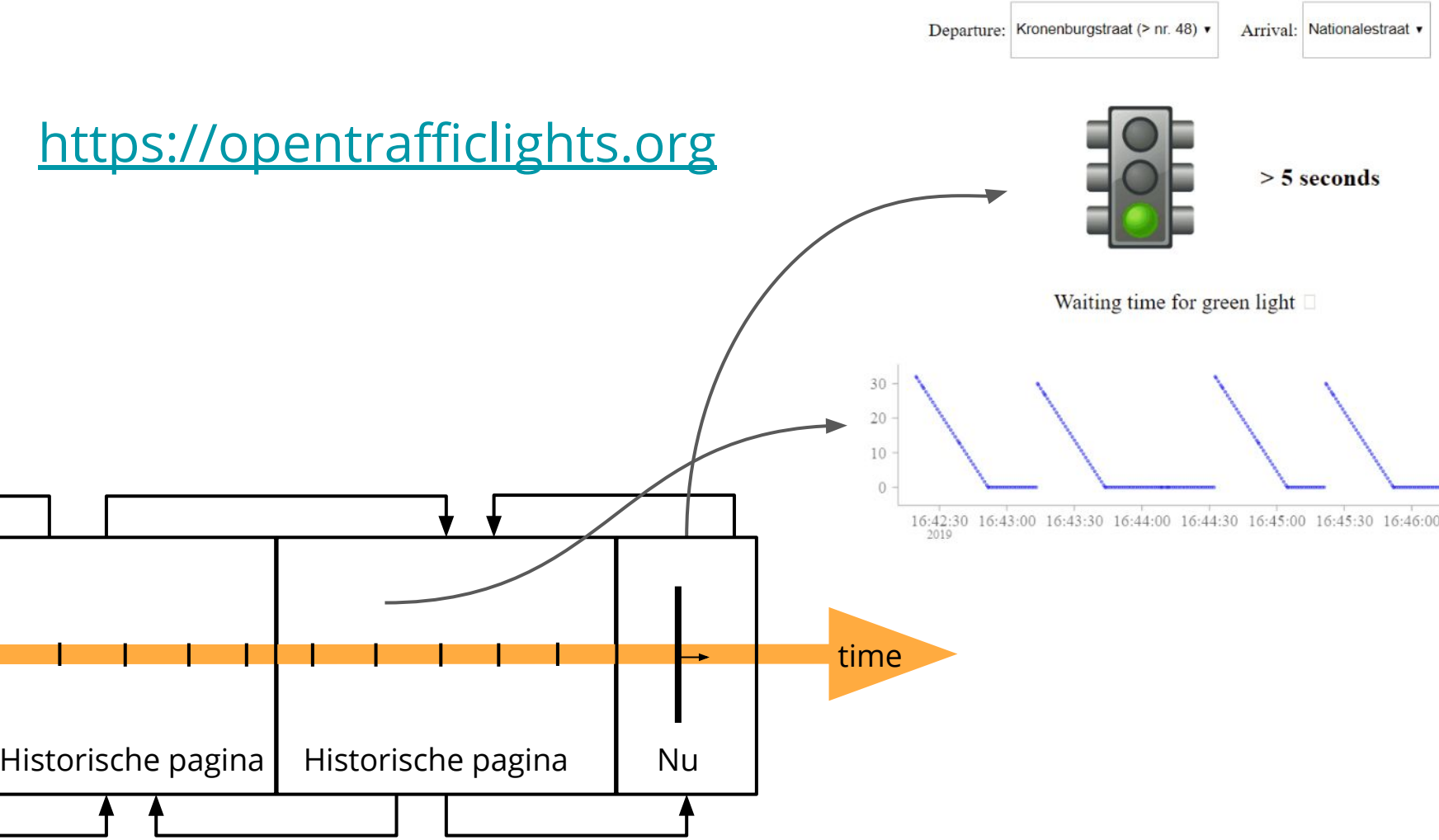


Open Traffic Lights ontology: https://lov.linkeddata.es/dataset/lov/vocabs/otl

SPAT describes the signal phase and timing of a signal group



https://opentrafficlights.org

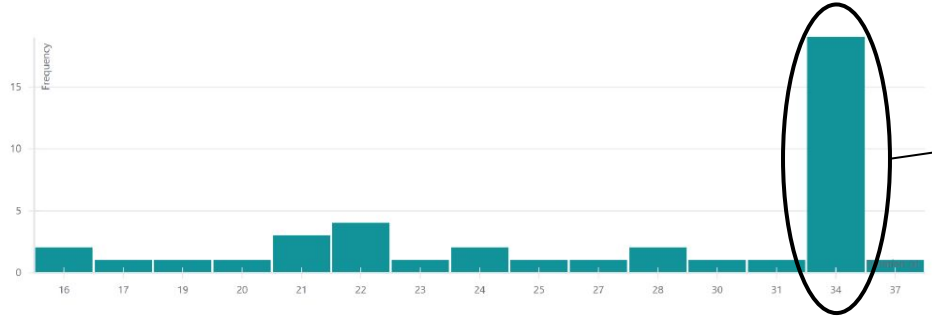


Semantics of traffic lights data

Predicting phase durations

Demo

1) Create frequency distributions of signal phase durations



Apparently, this phase takes mostly 34 seconds (for this signal group)

Historical dataset collected for 3 weeks

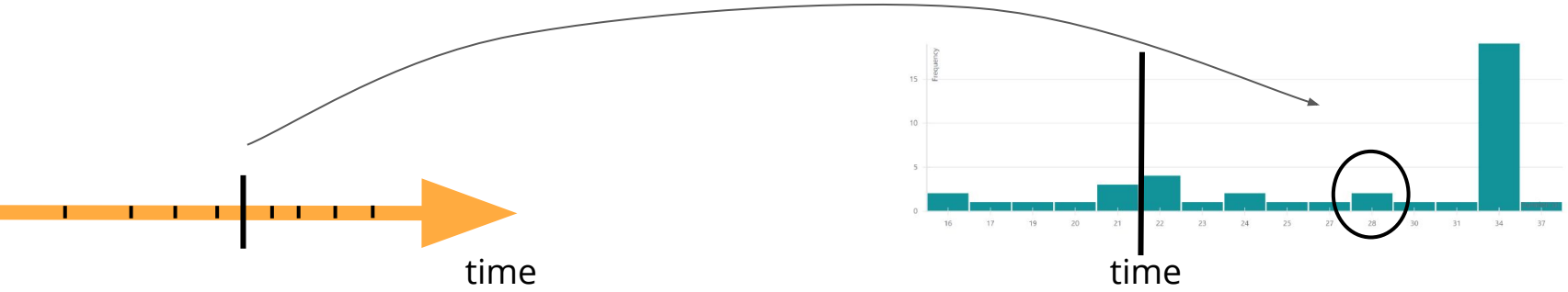
Bucket strategies:

+ Only per signal group and signal phase

+ Per type of day (weekday or weekend) and hour (cfr. Morning traffic on a weekday between 8 and 9 a.m.)

+ Per day (monday) and time slots of 20 minutes

2) Predict the current phase duration



Predicting duration for SPAT update i...

...with frequency distribution of other 9 groups

10-fold cross validation: dataset split in 10 groups of SPAT updates and corresponding frequency distributions

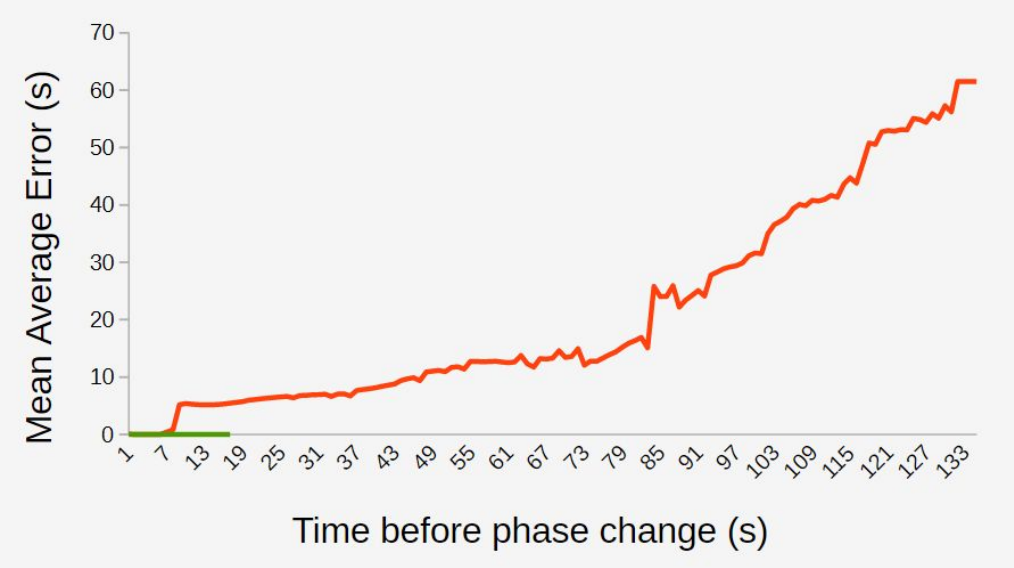
Mean average error $MAE = \frac{1}{n}\sum_{i=1}^{n}|d_p(i) - d(i)|$

Lowest MAE with median and fine-grained grouping

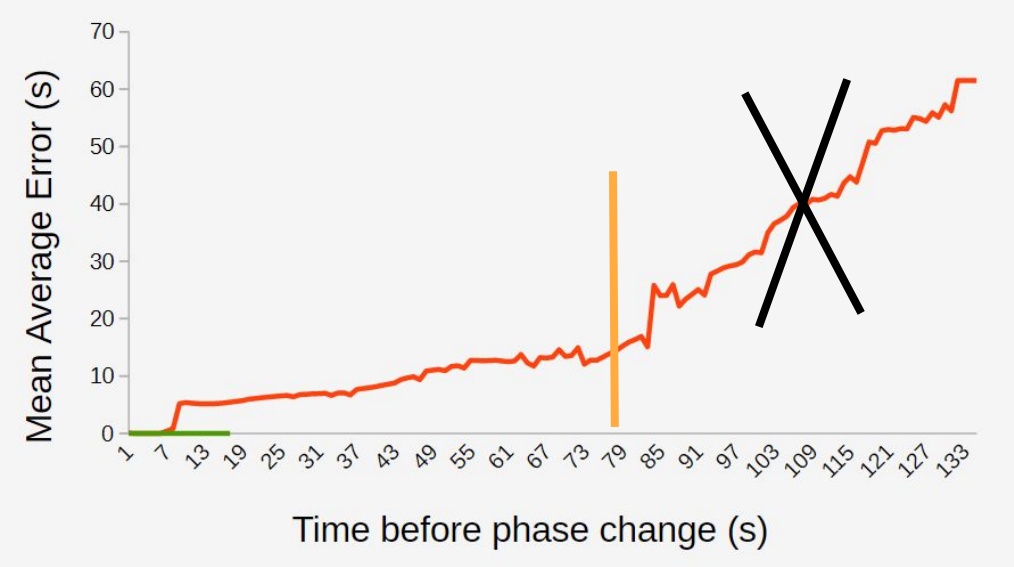
Method	No grouping (s)	Per type of day and every hour (s)	Per day and every 20 minutes (s)
Median	6.8	5.5	5.1
Mean	7.0	5.9	5.6
Mode	7.6	6.2	6.0

But what is the causality between the prediction error and the time before end of phase?

Green is more predictable than red and the prediction can be totally useless



Let users set a fixed probability as a user preference



E.g. 90% sure that phases will not take longer than 79 seconds

Discussion point 1

How can we add more context to lower the prediction error and detect exceptionally cases?

Publish live vehicle counter dataset

DCAT-AP, NGSI-LD context broker?

Semantics of traffic lights data

Predicting phase durations

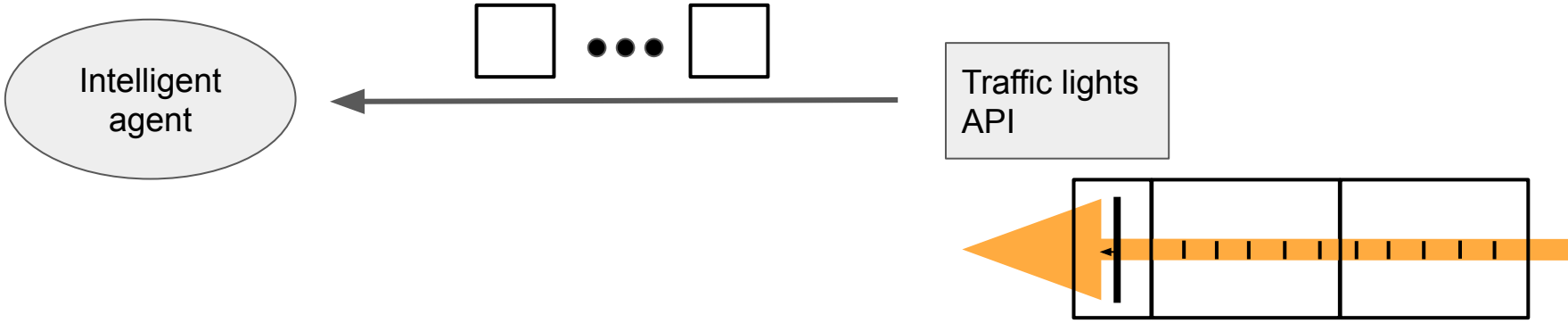
Demo

Client-side prediction of the phase duration

https://brechtvdv.github.io/Article-Predicting-traffic-light-phases/

Discussion point 2

When should we do statistics on the server-side or client-side?



Future work: semantically describe statistics (create summaries, average, median…) that can be published server-side or generated client-side

Predicting phase durations of traffic lights using live Open Traffic Lights data

9th September 2019
International Workshop On Semantics For Transport 2019

Brecht Van de Vyvere – Karel D'Haene - Kurt D'Haene - Pieter Colpaert – Ruben Verborgh
IDLab (IMEC – Ghent University)

imec

UNIVERSITEIT GENT