



A State-Space Model For Assimilating Passenger And Vehicle Flow Data With User Feedback In A Transit Network

Sylwester Arabas & Alexandros Papacharalampous
AETHON Engineering Consultants, Athens, Greece



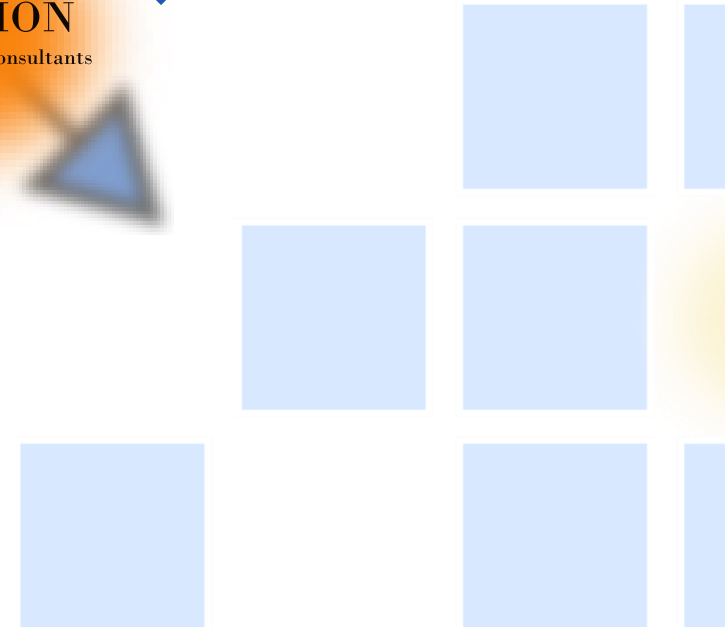
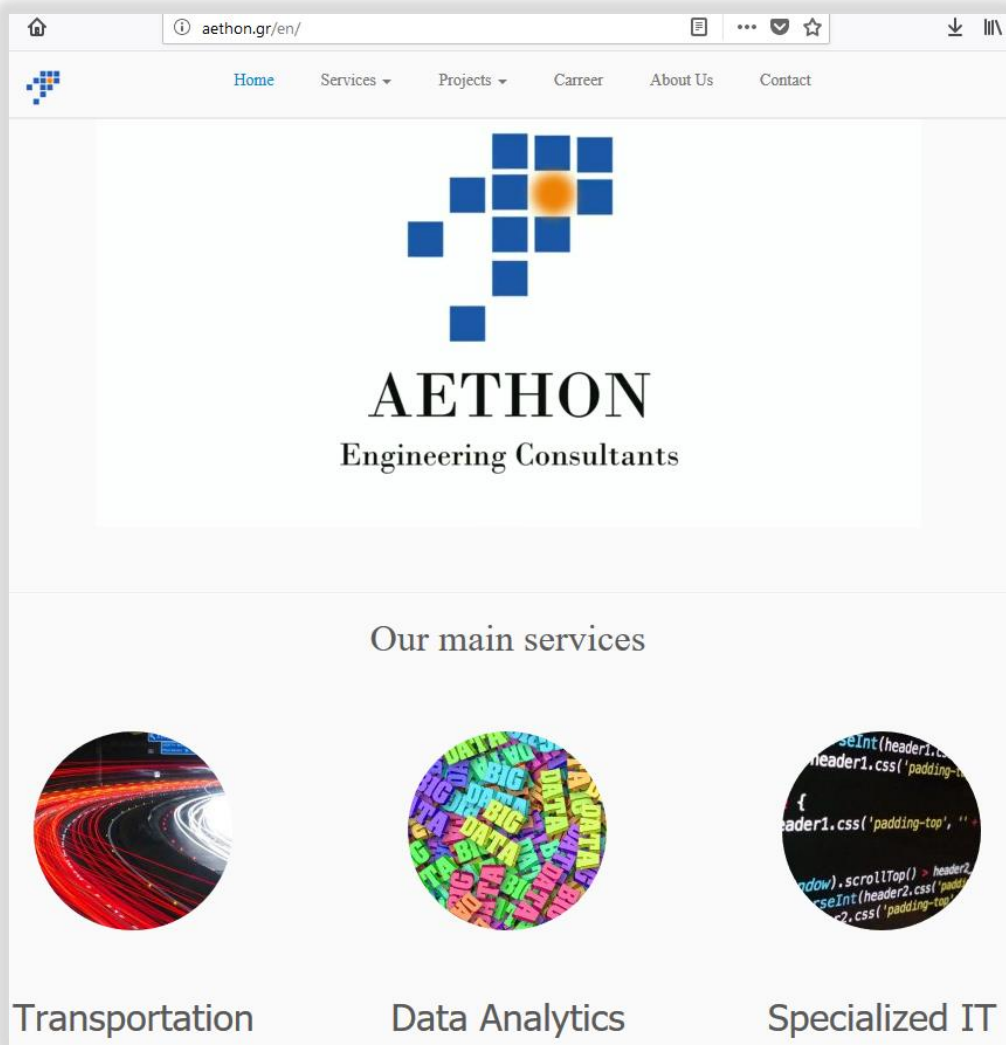
MATTS 2018, TUDelft, October 2018





Alexandros Papacharalampous

Sylwester
Arabas



idea!

Tour Eiffel - Parc du Champ-de-Mars, 75
 Cimetière-Ville de Levallois, 101 Rue Bai

Send directions to your phone

via Rue Victor Hugo/D909	26 min	6.7 km
Fastest route now due to traffic conditions		
via Bd Périphérique	31 min	7.6 km
Heavy traffic, as usual		
via Bd Périphérique and Rue Victor Hugo/D909	32 min	9.4 km
Lighter traffic than usual		

Tour Eiffel - Parc du Champ-de-Mars, 75
 Cimetière-Ville de Levallois, 101 Rue Bai

3:45 PM–4:36 PM	51 min
3:59 PM from École Militaire €1.90 24 min every 3 min	
3:45 PM–4:30 PM 45 min	
3:50 PM–4:38 PM 48 min	
3:45 PM–4:30 PM 45 min	

solution?

Why one would care about crowdedness?

see e.g. "Crowding in public transport: Who cares and why?": Haywood et al. 2017

- ability to seat/work/board (uncertainty)
- safety, security, hygiene, thermal comfort (stress)
- infrastructure/resource use (inefficiencies)

How and what for to use information on crowdedness?

see e.g. "A Mesoscopic Transit Assignment Model Including Real-Time Predictive Information on Crowding": Nuzzolo et al. 2016,
"Simulating the effects of real-time crowding information in public transport networks": Drabicki et al. 2017,
"Impact of real-time crowding information: a Stockholm metro pilot study": Zhang et al. 2017

- route planning (route/mode choice)
- real-time information provision (load optimisation)
- planning ... (operator's perspective)

What data could one use (assimilate)?

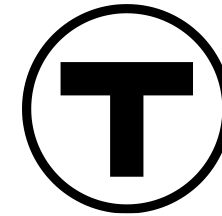
- vehicle positions
- schedules / headway times
- vehicle weighing (e.g., Nielsen et al. 2014, Jenelius 2018)
- station gate entry/exit counts
- smartcard data
- automatic passenger counters (APCs) (e.g., Pinna & Dalla Chiara 2010)
- image recognition (e.g., Toyosawa & Kawai 2005)
- comm. device counting - WiFi/GSM (e.g., Schauer et al 2014)
- user feedback (e.g., Moovit)



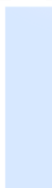
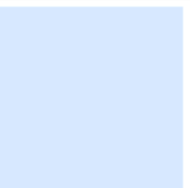
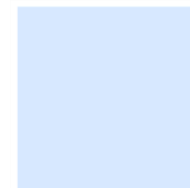
simple state-space model?

$$\begin{aligned}x_k &= F_k x_{k-1} + B_k u_k + w_k \\z_k &= H x_k + v_k\end{aligned}$$

minimal test-case dataset?



prototype implementation?

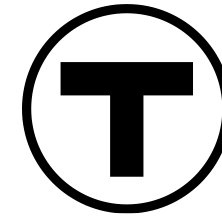




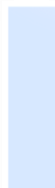
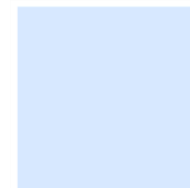
simple state-space model?

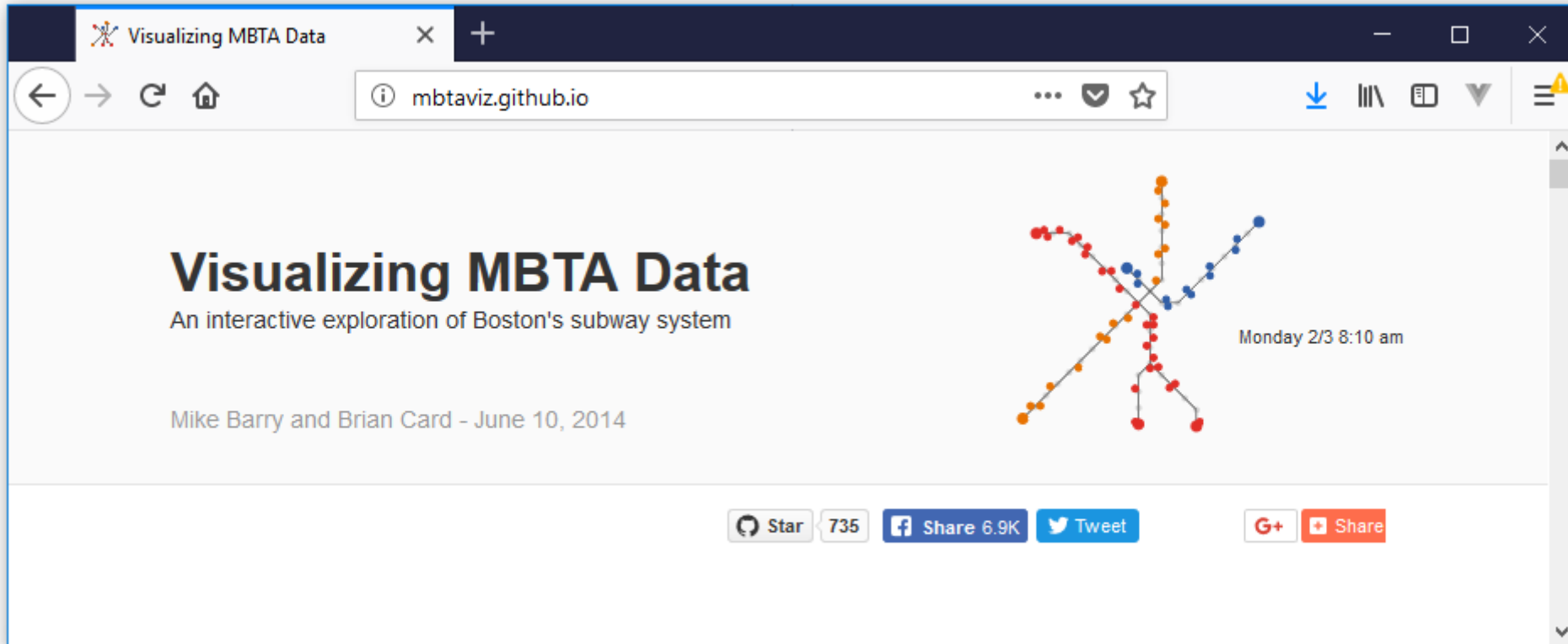
$$\begin{aligned}x_k &= F_k x_{k-1} + B_k u_k + w_k \\z_k &= H x_k + v_k\end{aligned}$$

minimal test-case dataset?



prototype implementation?

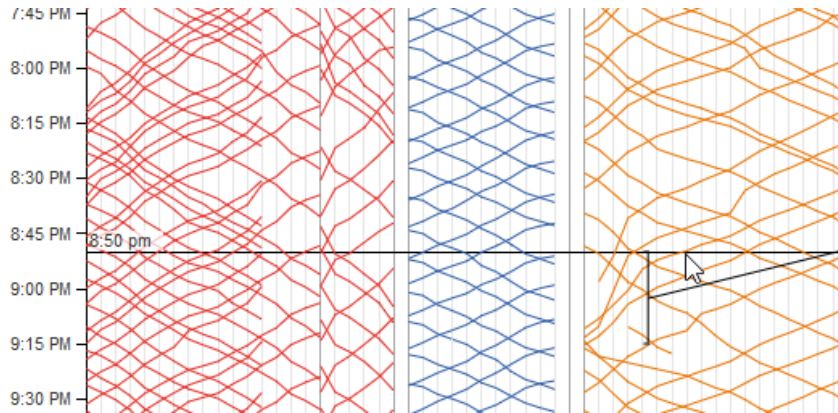
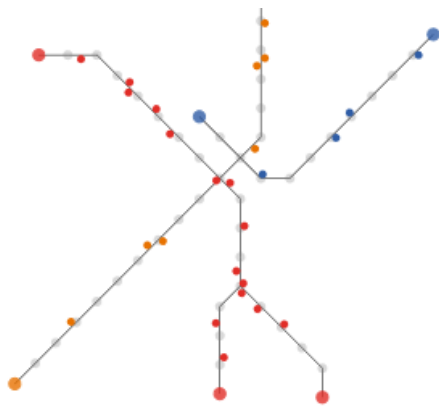
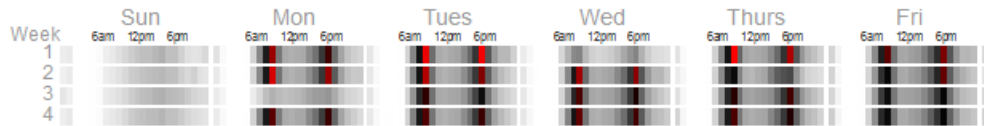




Accompanied by an open dataset released by Boston's Massachusetts Bay Transit Authority (MBTA):

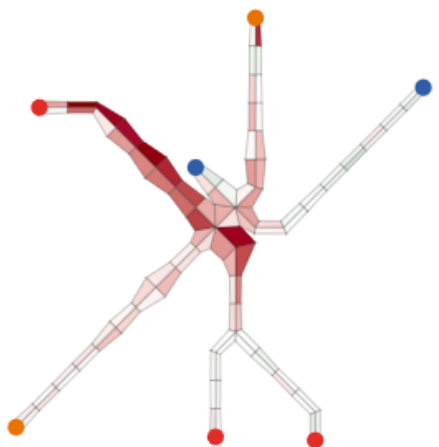
- 3 subway lines, 4 weeks of data
- turnstile counts with 1-min. time resolution
- train position data
- archive of alerts

Entrances and Exits from All Stations during February 2014

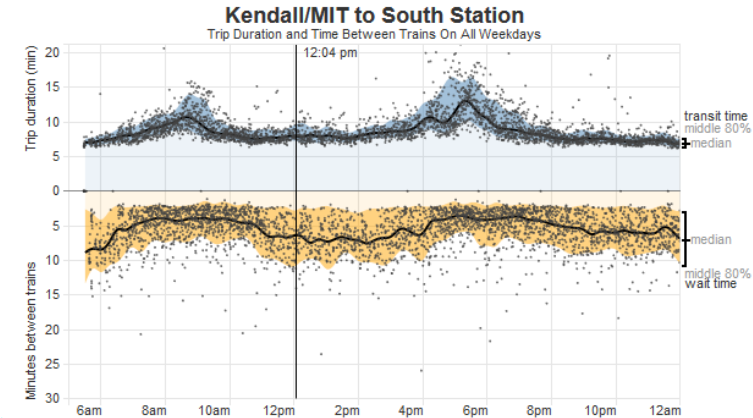
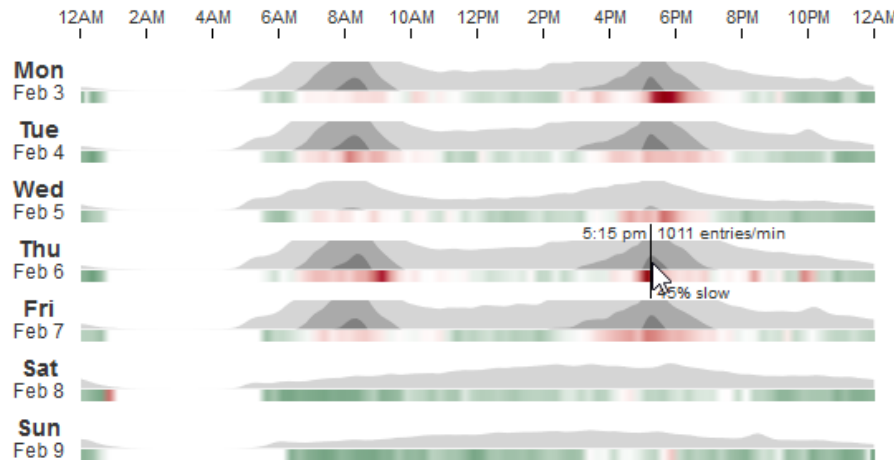


A disabled train at Wellington Station causes northbound delays on the Orange Line from 8:50PM to 9:15PM

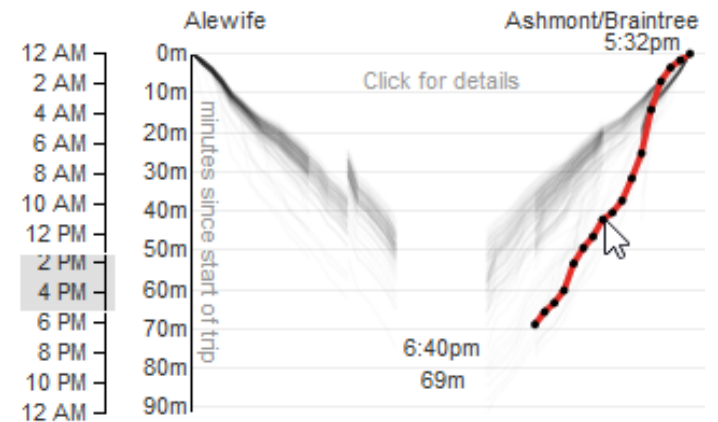
Notice how southbound trains are temporarily delayed, but get back on schedule quickly.



5:15 pm on Thu Feb 6



At 12:04 pm trains leave every 3 to 11 minutes from Kendall/MIT going to South Station. The trip takes between 7 and 10 minutes. The shortest time from when you walk into Kendall/MIT until you walk out of South Station is 7 minutes but it can be as long as 21 minutes. Usually it takes about 11 minutes including wait and transit time.



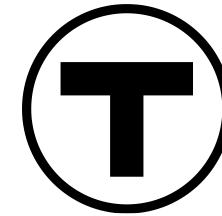


simplest state-space model?

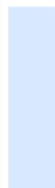
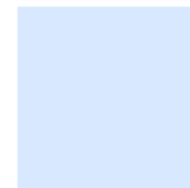
$$\begin{aligned}x_k &= F_k x_{k-1} + B_k u_k + w_k \\z_k &= H x_k + v_k\end{aligned}$$



minimal test-case dataset?



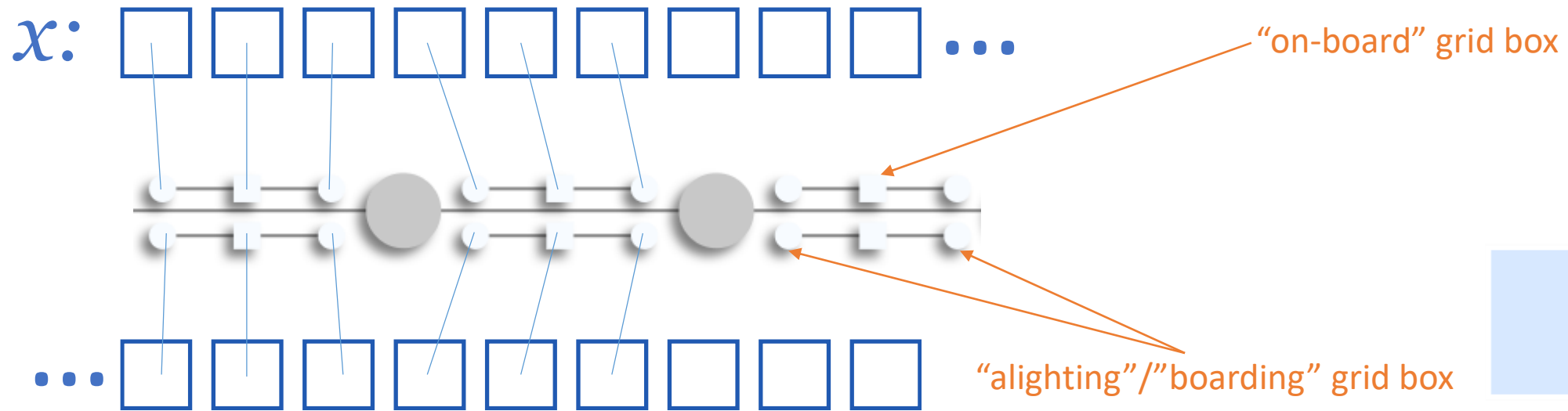
prototype implementation?



state vector x
 (k : time level)

$$x_k = F_k x_{k-1} + B_k u_k + w_k$$

$$z_k = H x_k + v_k$$



state transition matrix F
 (k : time level)

$$x_k = F_k x_{k-1} + B_k u_k + w_k$$

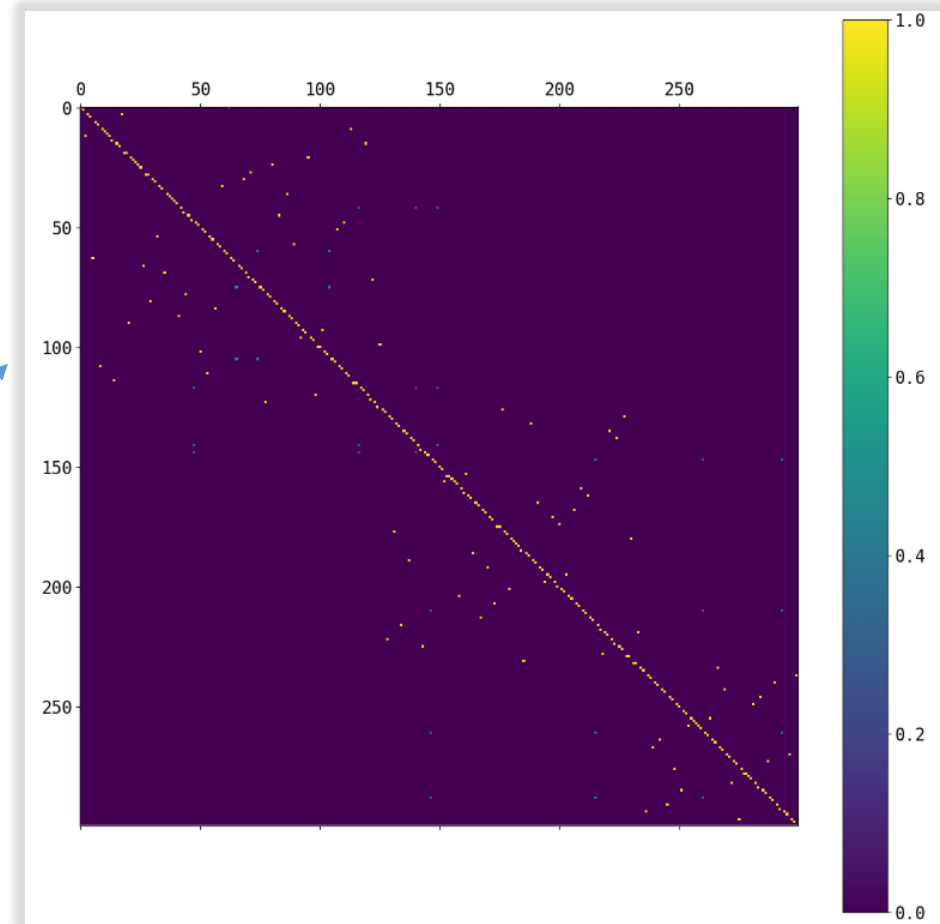
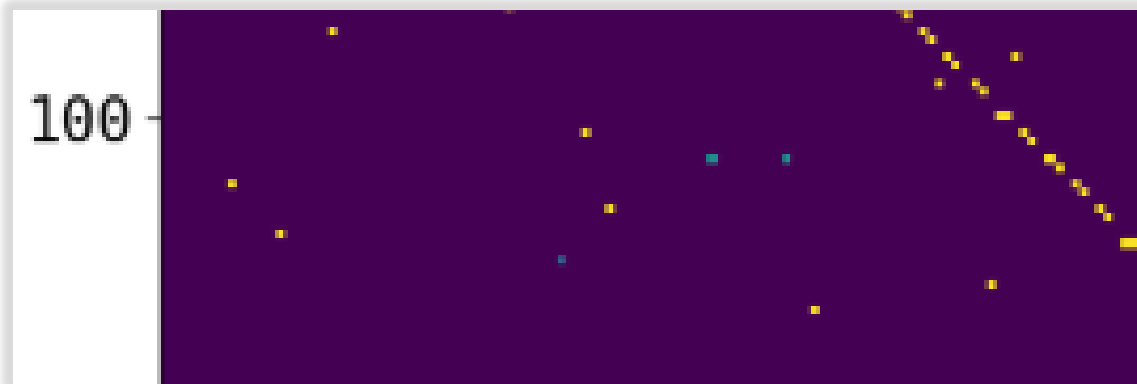
$$z_k = H x_k + v_k$$

single-line case:

F : Boolean bidiagonal time-dependent matrix

multi-line network:

F includes probabilities of transfer
 (rationale for alighting/boarding grid boxes)



control vector u
 (k : time level)

$$x_k = F_k x_{k-1} + B_k u_k + w_k$$

$$z_k = H x_k + v_k$$



$$n_u \neq n_x$$

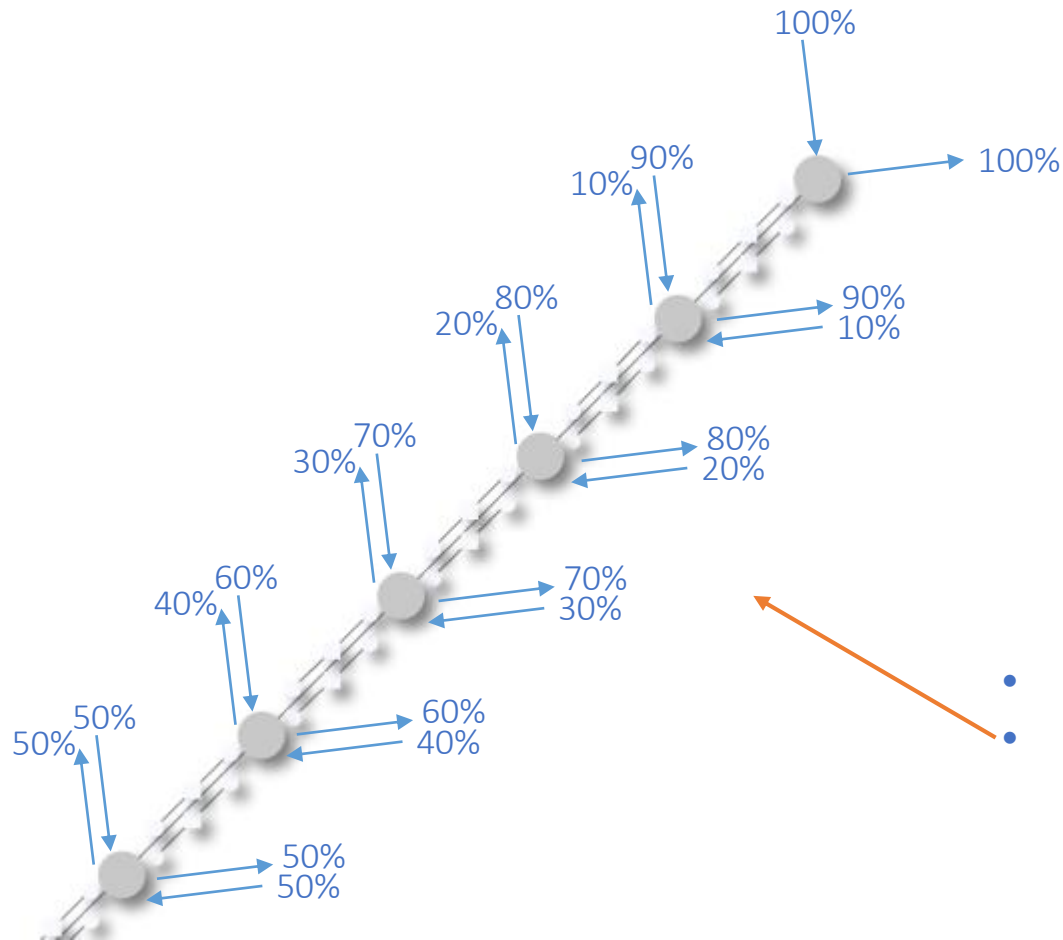
NB: interplay between Fx and Bu makes the model capable of:

- representing accumulation of waiting passengers on platforms
- coping with (slightly) unsynced passenger/vehicle data

control transition matrix B
 (k : time level)

$$x_k = F_k x_{k-1} + B_k u_k + w_k$$

$$z_k = H x_k + v_k$$



NB: conservation of total passenger count (constraint on B and F):

$$\sum u_k + \sum x_{k-1} = \sum x_k$$

- OD matrix-based?
- simpler: constant B ?

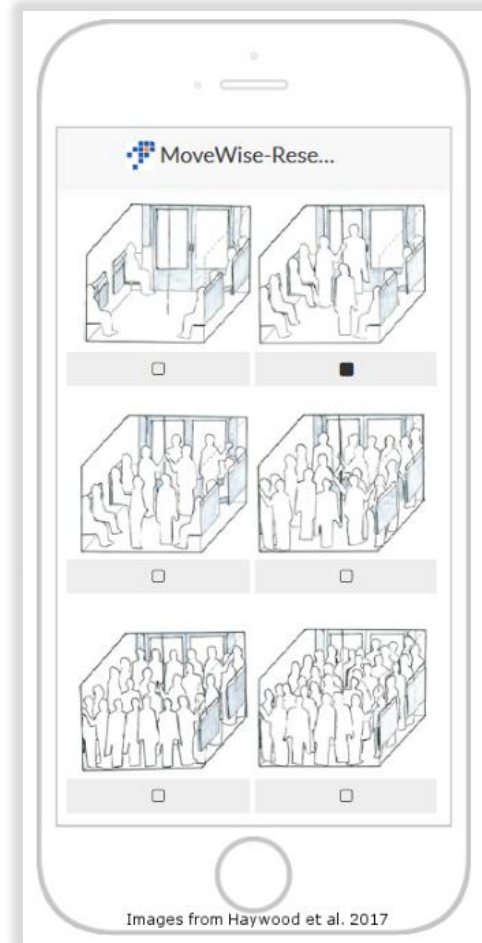


measurement vector \mathbf{z}
 observation matrix \mathbf{H}

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{B}_k \mathbf{u}_k + \mathbf{w}_k$$

$$\mathbf{z}_k = \mathbf{H} \mathbf{x}_k + \mathbf{v}_k$$

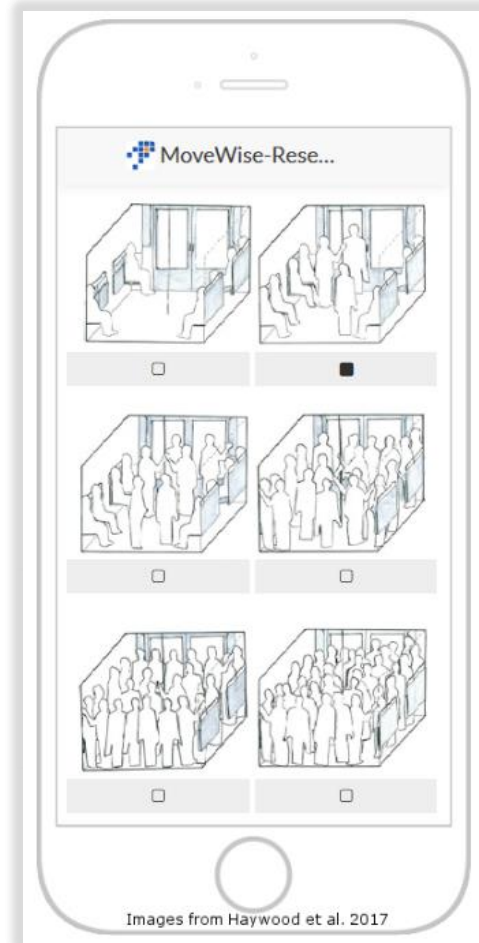
- unlike the $\mathbf{B}\mathbf{u}$ term, \mathbf{z} and \mathbf{H} are applicable to integral measurements (differential data corresponding to entry/exit counts)
- \mathbf{H} allows to express spatial granularity assumptions
- not limited to one measurement type!



Gaussian noise terms w, v (k : time level)

$$x_k = F_k x_{k-1} + B_k u_k + w_k$$

$$z_k = H x_k + v_k$$



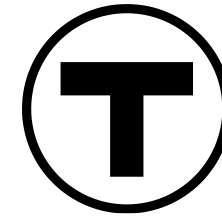
- sources of uncertainty: model & measurement
- prospect for applying Kalman filtering (then, different uncertainties for different measurement types)



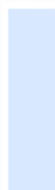
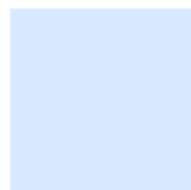
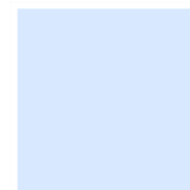
simple state-space model?

$$\begin{aligned}x_k &= F_k x_{k-1} + B_k u_k + w_k \\z_k &= H x_k + v_k\end{aligned}$$

minimal test-case dataset?



prototype implementation?





MoveWise prototype

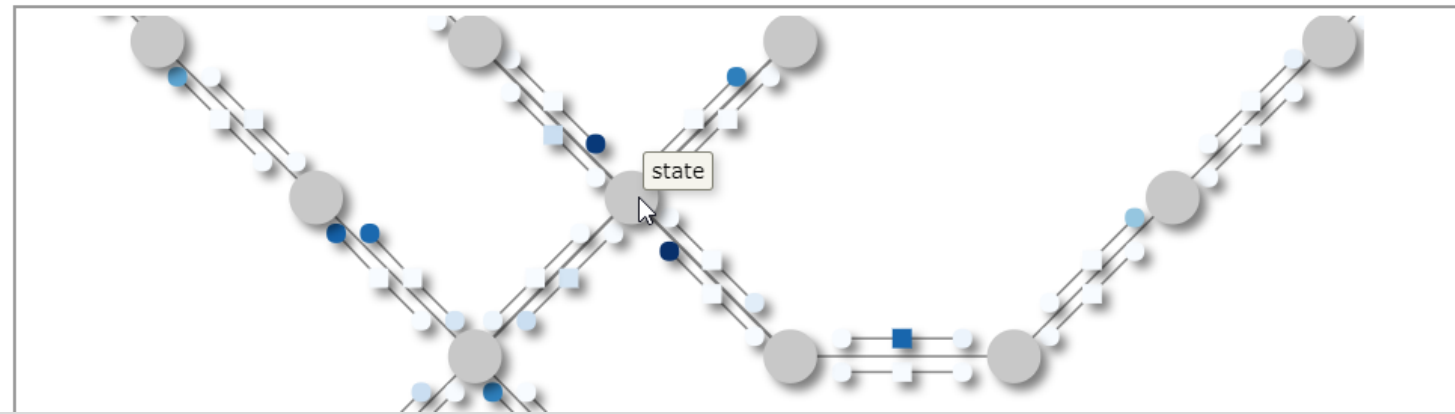
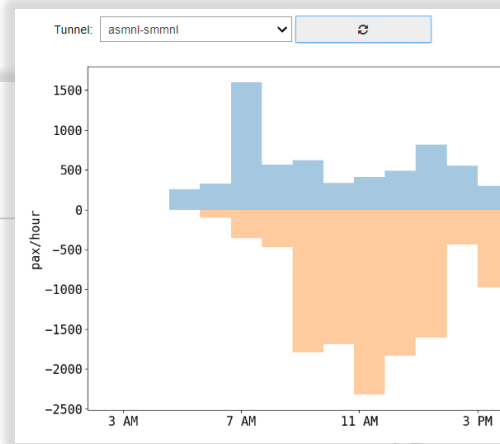
Mon Feb 17 15:42:00 2014

23805

Alert: 2014-02-13 20:45:00-05:00 - 2014-02-18 20:45:00-05:00

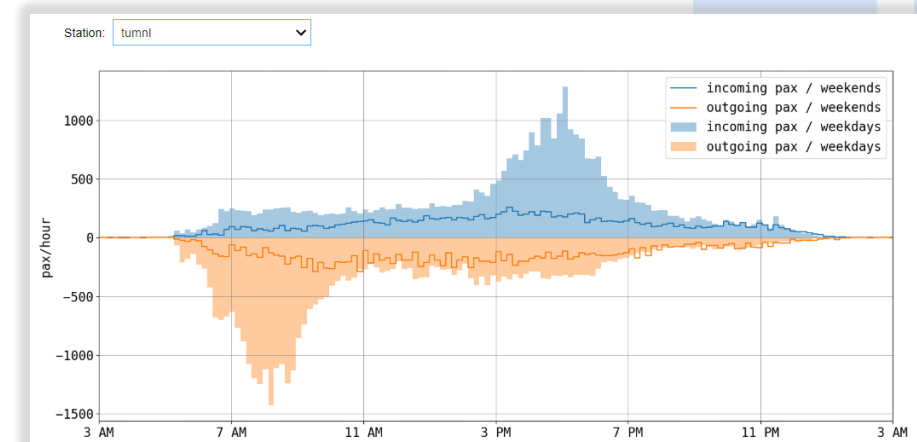
Feb 13, 18, 24-25: Shuttle buses replacing Orange Line trains between Sullivan Square and Oak Grove from approx 8:45 p.m. to end of service

Navigation: -24h, -1h, +1h, +24h

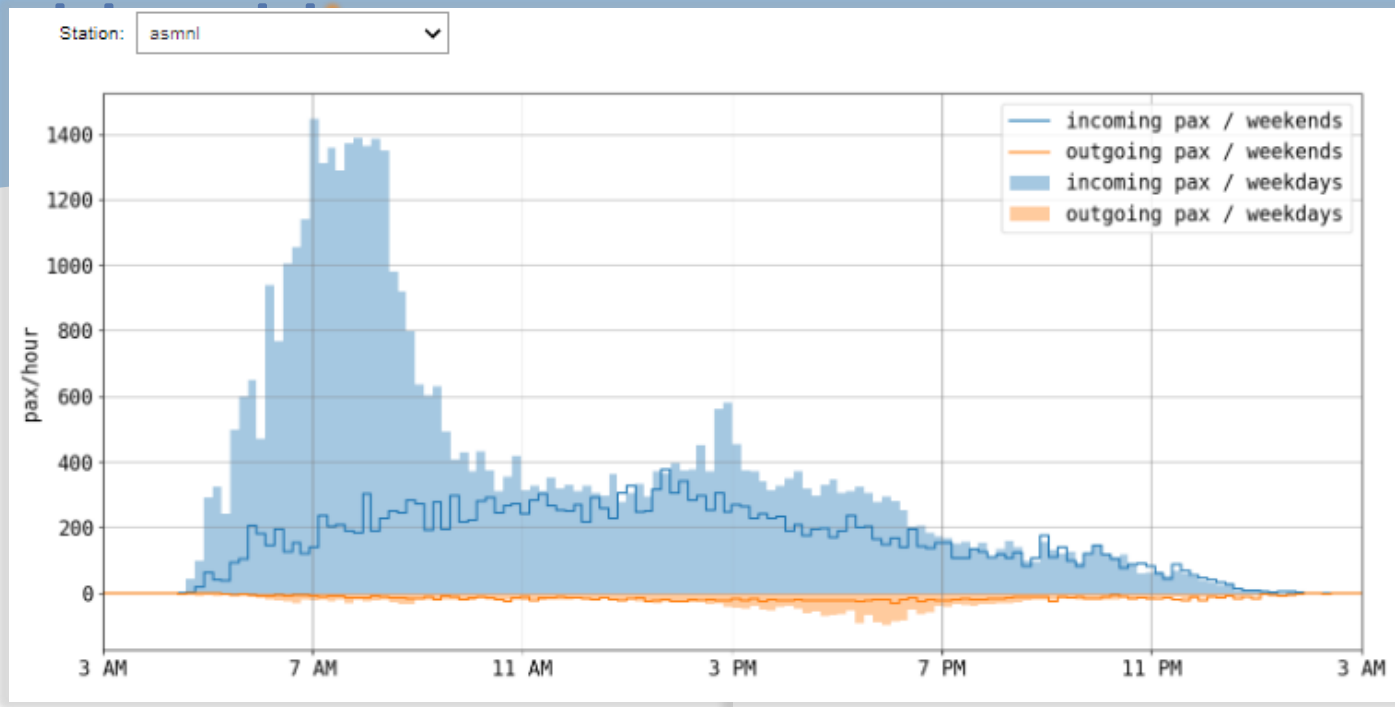


turnstile data plots | tunnel plots | feedback emulation | network map | matrix viewer

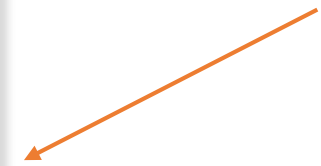
Scope: haecI-north-tnl, north-haec-tnl, jaksn-rcmnl-tnl, rcmnl-jaksn-tnl, masta-bbsta-tnl, bbsta-masta-tnl, mimnl-ogmnl-tnl, ogmnl-mimnl-tnl, north-ccmnl-tnl, ccmnl-north-tnl, rugg-rcmnl-tnl, rugg-masta-tnl, masta-rugg-tnl, sbmnl-jaksn-tnl, jaksn-sbmnl-tnl, state-haec-tnl, haecI-state-tnl, sull-welin-tnl, welin-sull-tnl, tumnl-chncI-tnl, chncI-tumnl-tnl, welin-mimnl-tnl, mimnl-welin-tnl, aport-wimnl-tnl, wimnl-apor-tnl, aqucl-mvbcI-tnl, mvbcI-aqucl-tnl, bmmnl-rbmnI-tnl, rbmnI-bmmnl-tnl, bomnl-gover-tnl, gover-bomnl-tnl, gover-state-tnl, state-gover-tnl, mvbcI-apor-tnl, apor-mvbcI-tnl, orhte-sdmnl-tnl, sdmnl-orhte-tnl, rbmnI-wondI-tnl, wondI-rbmnI-tnl, sdmnl-bmmnl-tnl, bmmnl-sdmnl-tnl, state-aqucl-tnl, aqucl-state-tnl, wimnl-orhte-tnl, orhte-wimnl-tnl



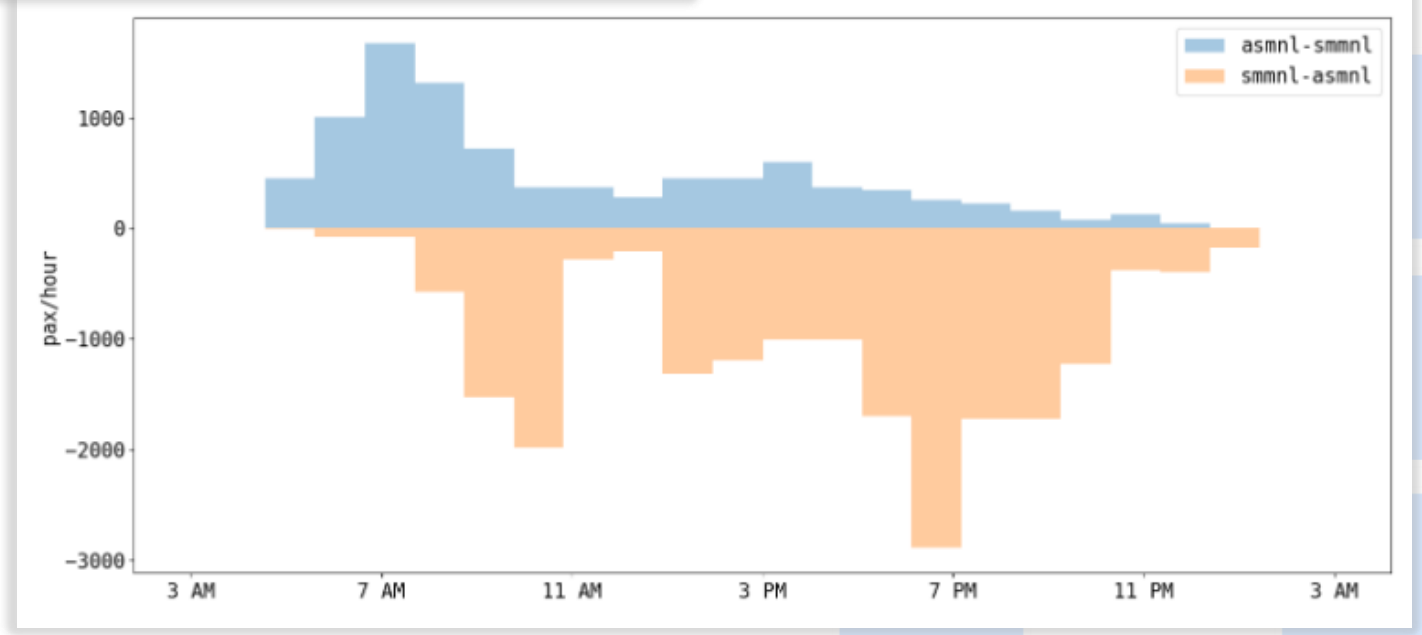
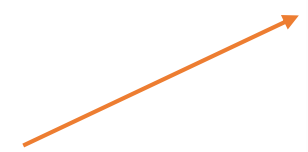
idea: use the state-space model and filter error covariance matrix for triggering feedback requests



input data (4-week average)
(outgoing pax not counted with turnstiles)



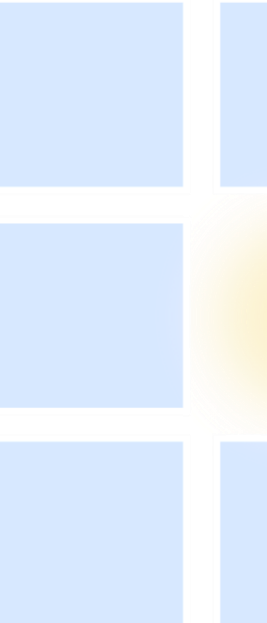
simulation output: Feb 25 2014
(on-board pax reaching terminus)



Summary:

$$x_k = F_k x_{k-1} + B_k u_k + w_k$$

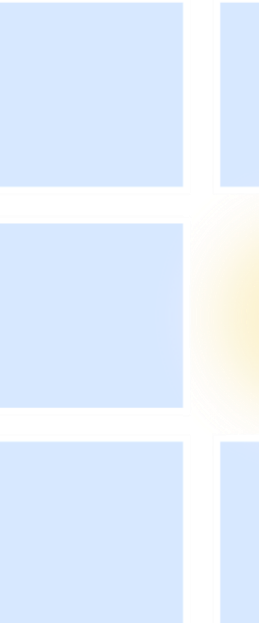
- linear model of the dynamics of passenger loading in a transit network congruent with the underlying eqs. of the Kalman filter with control input



Summary:


$$x_k = F_k x_{k-1} + B_k u_k + w_k$$

- linear model of the dynamics of passenger loading in a transit network congruent with the underlying eqs. of the Kalman filter with control input
- data fusion methodology for passenger and vehicle flows exemplified with the MTBAViz dataset



Summary:

$$x_k = F_k x_{k-1} + B_k u_k + w_k$$

- linear model of the dynamics of passenger loading in a transit network congruent with the underlying eqs. of the Kalman filter with control input
- data fusion methodology for passenger and vehicle flows exemplified with the MTBAViz dataset 
- data assimilation potential for vehicle weighing, APC, WiFi/GSM or real-time user feedback-based measurements

Thank you for your attention!

free & open-source software used in implementation:

Jupyter, Python, numpy,
ipytest, ipywidgets, visJS2jupyter,
networkX, filterpy



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 739607.

