Timetabling for Passengers: A Knock-On Delay Model

Peter Sels[1,2,3], Thijs Dewilde[1], Dirk Cattrysse[1], Pieter Vansteenwegen[1]

[1]Katholieke Universiteit Leuven,
Centre for Industrial Management/Traffic & Infrastructure,
Celestijnenlaan 300a, 3001 Heverlee, Belgium

[2]Logically Yours BVBA,
Plankenbergstraat 112 bus L7, 2100 Antwerp, Belgium
e-mail: sels.peter@gmail.com, corresponding author

[3]Infrabel, Department of Rail Access,
Frankrijkstraat 91, 1070 Brussels, Belgium

December 5, 2013

Table of Contents

Task

Belgian Infrastructure Management Company: Infrabel:

Add Knock-on Delays as a term to
Expected Passenger Travel Time Goal Function

Goals:

Reduce Expected Passenger Time $\Rightarrow$ Optimises Robustness

Fixed:

Infrastructure, Train Lines, Halting Pattern, Primary Delay Distributions

Variable:

Timing: Supplement Times at every Ride, Dwell, Transfer Action,
$\Rightarrow$ variable inter-Train Heading Times $\Rightarrow$ variable Train Orders

Specifics:

One Busy Day, Morning Peak Hour

Context: FAPESP: Two Phased

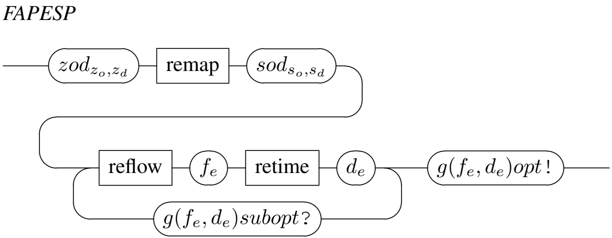
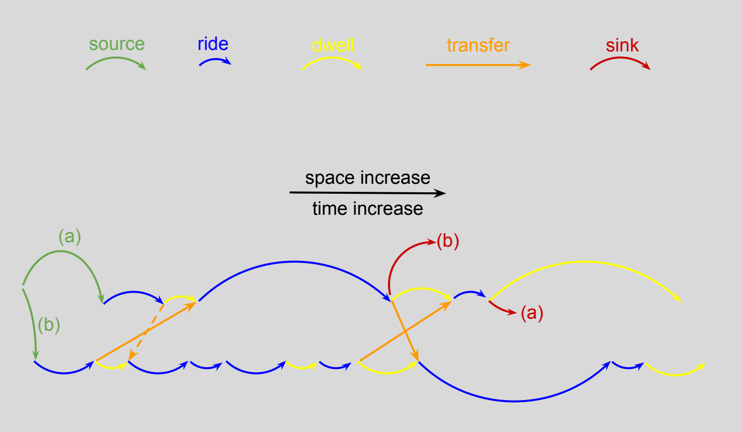
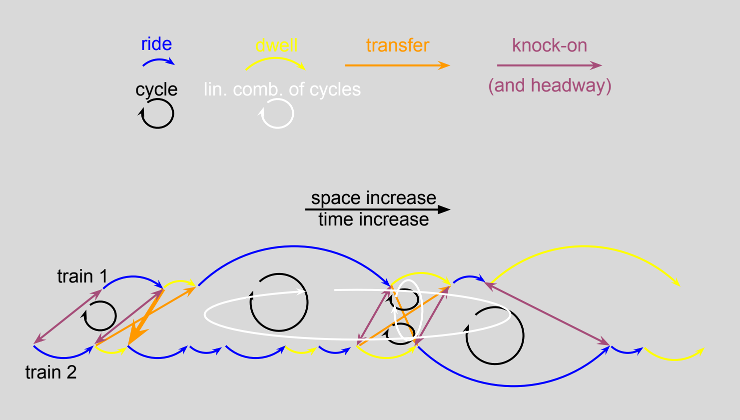


Figure: Two Phased implies Iterations

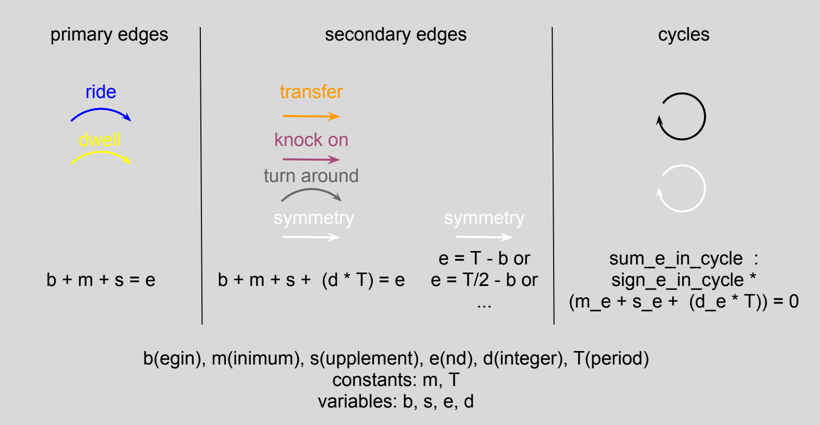
Graph for Reflowing: add Source & Sink Edges



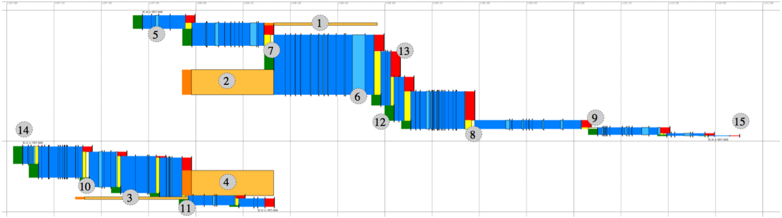
Graph for Retiming: add Knock-On Edges & Cycles



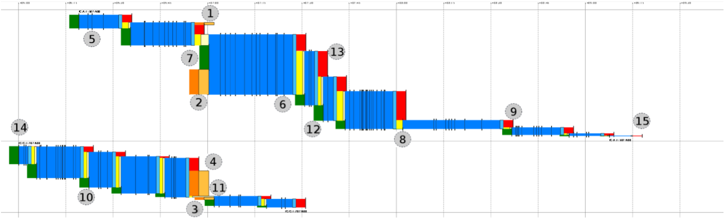
Graph for Retiming: All Constraints



Reflowing decides on Rectangle Heights
Retime (=Timetabling) decides on Rectangle Widths

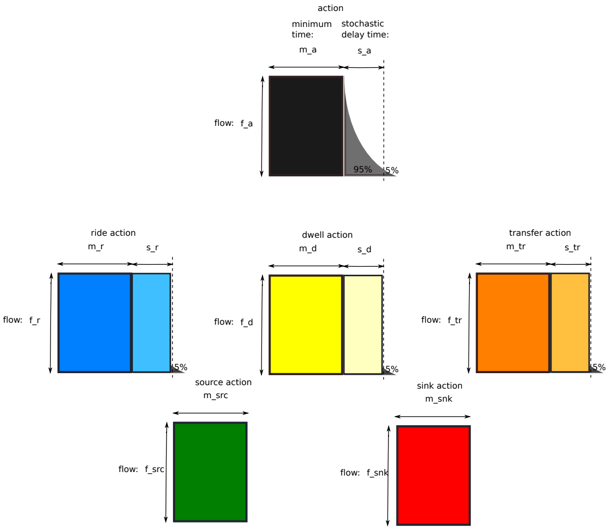


(a) Original Schedule



(b) Optimized Version

Action: Negative Exponential Delay Distribution



Stochastic Goal Function: Expected Passenger *Transfer* Time

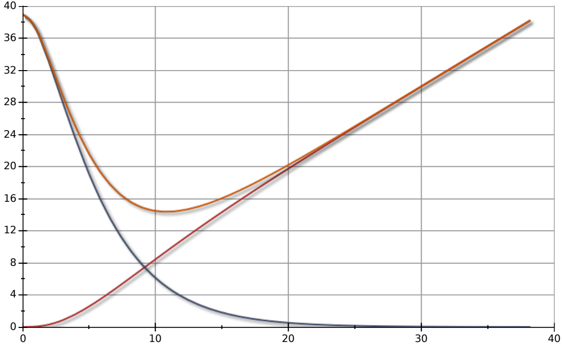
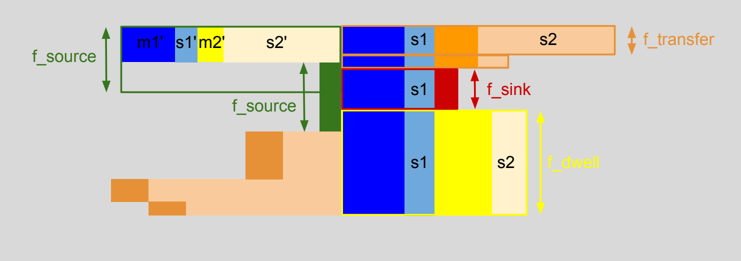


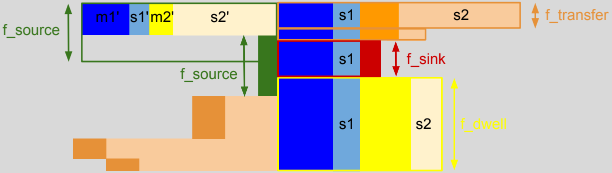
Figure: D_0 is introduced supplement, $D_1 > D_0$ is delta time of next chance action. Curve maps planned time to expected time.

Grouping per Subsequent Action-Pair

- departing = ride' + dwell' + source
- through = ride + dwell
- changing = ride + transfer
- arriving = ride + sink



Grouping per Subsequent Action-Pair towards Cost

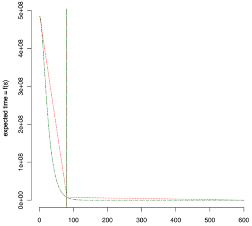


flows	domain = planning: planned time	domain = execution: cost = expected time
f_departing = f_source	m1' + s1' + m2'+ s2'	m1' + m2'+ cost(s1'+ s2')
f_through = f_dwell	m1 + s1 + m2+ s2	m1 + m2+ cost(s1+ s2)
f_changing = f_transfer	m1 + s1 + m2+ s2	m1 + m2 + cost(s1+ s2)
f_arriving = f_sink	m1 + s1	m1 + cost(s1)

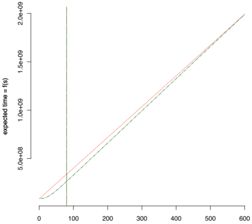
In-Time and Over-Time

	In-Time	Over-Time
probability	$\int_0^{D_0} p_a(x)dx$	$\int_{D_0}^{D_1} p_a(x)dx$
inc./dec. in D_0	inc.	dec.
expected time	$\int_0^{D_0} p_a(x)D_0dx$	$\int_{D_0}^{D_1} p_a(x)D_1dx$
inc./dec. in D_0	inc.	dec.
departing = ride' + dwell' + source		✓
through = ride + dwell	✓	
changing = ride + transfer	✓	✓
arriving = ride + sink	✓	

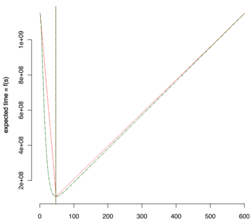
Cost curves of 4 Passenger Categories



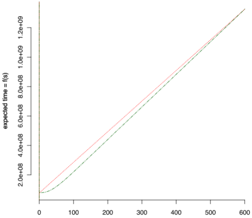
(a) departing=ride'+dwell' +source



(b) through=ride+dwell



(c) changing=ride+transfer

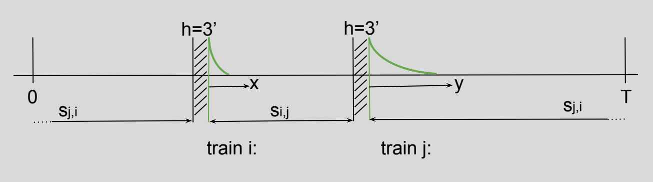


(d) arriving=ride+sink

Primary Delay Distributions

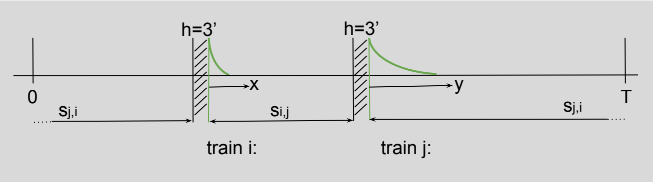
$$p_i(x) = a_i e^{-a_i x}, p_j(y) = a_j e^{-a_j y}, \quad (1)$$

$$\overline{c_i} = \int_0^\infty x a_i e^{-a_i x} dx = \frac{1}{a_i}, \overline{c_j} = \int_0^\infty y a_j e^{-a_j y} dy = \frac{1}{a_j}. \quad (2)$$



Knock-On Probability Derivation

Probability of knock-on delay



Integrate over 2 triangle areas where the delay difference

- $x \geq y + s_{i,j}$
- $y \geq x + s_{j,i}$

as in

$$\begin{aligned} p_{x \geq y + s_{i,j}}(a_i, a_j, s_{i,j}) &= \int_0^\infty \int_{y+s_{i,j}}^\infty a_i e^{-a_i x} \cdot a_j e^{-a_j y} \, dxdy = \frac{a_j e^{-a_i s_{i,j}}}{a_i + a_j}, \\ p_{y \geq x + s_{i,j}}(a_i, a_j, s_{j,i}) &= \int_0^\infty \int_{x+s_{j,i}}^\infty a_i e^{-a_i x} \cdot a_j e^{-a_j y} \, dydx = \frac{a_i e^{-a_j s_{j,i}}}{a_i + a_j}. \end{aligned} \tag{3}$$

Knock-On (Train & Passenger) Time Derivation

Train Time Cost of knock-on delay

$$\begin{aligned} tKO_{i,j}(a_i, a_j, s_{i,j}) &= \int_0^\infty \int_{y+s_{i,j}}^\infty \underbrace{a_i e^{-a_i x} \cdot a_j e^{-a_j y}}_{probabilty} \underbrace{(x - y - s_{i,j})}_{tko_{i,j} \geq 0} dxdy \\ &= \frac{a_j e^{-a_i s_{i,j}}}{a_i(a_i + a_j)}, \\ tKO_{j,i}(a_i, a_j, s_{j,i}) &= \int_0^\infty \int_{x+s_{j,i}}^\infty \underbrace{a_i e^{-a_i x} \cdot a_j e^{-a_j y}}_{probability} \underbrace{(y - x - s_{j,i})}_{tko_{j,i} \geq 0} dydx \\ &= \frac{a_i e^{-a_j s_{j,i}}}{a_j(a_i + a_j)}. \end{aligned} \tag{4}$$

Passenger Time Cost of knock-on delay

$$\begin{aligned} pKO_{i,j}(a_i, a_j, s_{i,j}) &= f_j \cdot tKO_{i,j} = f_j \cdot \frac{a_j e^{-a_i s_{i,j}}}{a_i(a_i + a_j)}, \\ pKO_{j,i}(a_i, a_j, s_{j,i}) &= f_i \cdot tKO_{j,i} = f_i \cdot \frac{a_i e^{-a_j s_{j,i}}}{a_j(a_i + a_j)}. \end{aligned} \tag{5}$$

Two Train Example: KO Formulas

$$h + s_{i,j} + h + s_{j,i} = T \text{ or equivalently } s_{j,i} = T - 2h - s_{i,j}. \quad (6)$$

$$\begin{aligned}
0 &= \frac{d}{ds_{i,j}}\left(pKO_{i,j} + pKO_{j,i}\right) \\
\Leftrightarrow 0 &= \frac{d}{ds_{i,j}}\left(f_j \cdot \frac{a_j e^{-a_i s_{i,j}}}{a_i(a_i+a_j)} + f_i \cdot \frac{a_i e^{-a_j(T-2h-s_{i,j})}}{a_j(a_i+a_j)}\right) \\
\Leftrightarrow 0 &= -f_j \cdot \frac{a_j e^{-a_i s_{i,j}}}{a_i+a_j} + f_i \cdot \frac{a_i e^{-a_j(T-2h-s_{i,j})}}{a_i+a_j} \\
\Leftrightarrow f_j \cdot a_j e^{-a_i s_{i,j}} &= f_i \cdot a_i e^{-a_j(T-2h-s_{i,j})} \\
\Leftrightarrow ln\left(\frac{f_j \cdot a_j}{f_i \cdot a_i}\right) &= -a_j(T - 2h - s_{i,j}) + a_i(s_{i,j}) \\
\Leftrightarrow s_{i,j} &= \frac{a_j(T-2h) + ln\left(\frac{f_j a_j}{f_i a_i}\right)}{a_i+a_j}
\end{aligned} \quad (7)$$

From symmetry:

$$s_{j,i} = \frac{a_i(T - 2h) + ln\left(\frac{f_i a_i}{f_j a_j}\right)}{a_i + a_j}. \quad (8)$$

Two Train Example: Supplement Calculation

Two trains with:

- train i: expected delay of $1/a_i = 3$ minutes and $f_i = 100$ passengers
- train j: expected delay of $1/a_j = 1$ minute and $f_j = 300$ passengers
- $T = 60$ minutes, period
- $h = 3$ minutes, headway time

would be spread according to equations (7) and (8)

- $s_{i,j} = \frac{a_j(T-2h)+ln\left(\frac{f_j a_j}{f_i a_i}\right)}{a_i+a_j} = \frac{1(60-2\cdot3)+ln(300\cdot1/(100\cdot1/3))}{1/3+1} = 42.15$ min.
- $s_{j,i} = \frac{a_i(T-2h)+ln\left(\frac{f_i a_i}{f_j a_j}\right)}{a_i+a_j} = \frac{1/3(60-2\cdot3)+ln(100\cdot1/3/(300\cdot1))}{1/3+1} = 11.85$ min.

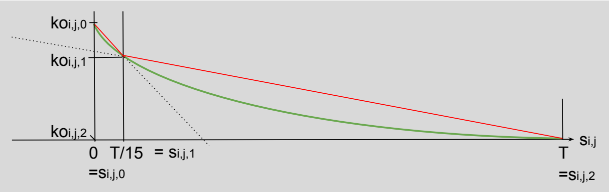
and indeed as equation (6) requires $42.15 + 3 + 11.85 + 3 = 60$ minutes.

All Knock-On Costs for $N(N-1)$ Trains on Same Resource: Formula

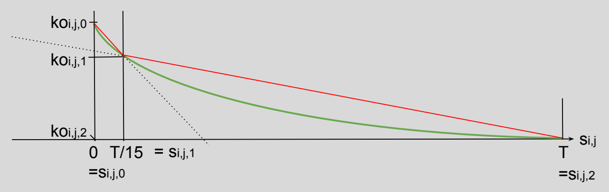
$$\forall R : pKO_R = \sum_{\substack{i,j \in I_R \\ i \neq j}} f_j \cdot \frac{a_j e^{-a_i s_{i,j}}}{a_i(a_i + a_j)}. \quad (9)$$

Is non-linear in $s_{i,j}$, but since we use convex minimisation $\Rightarrow$ use trick:

$$\forall R : \forall_{\substack{i,j \in I_R \\ i \neq j}} : pKO_{R,i,j} \geq f_j \cdot \frac{a_j e^{-a_i s_{i,j}}}{a_i(a_i + a_j)}. \quad (10)$$



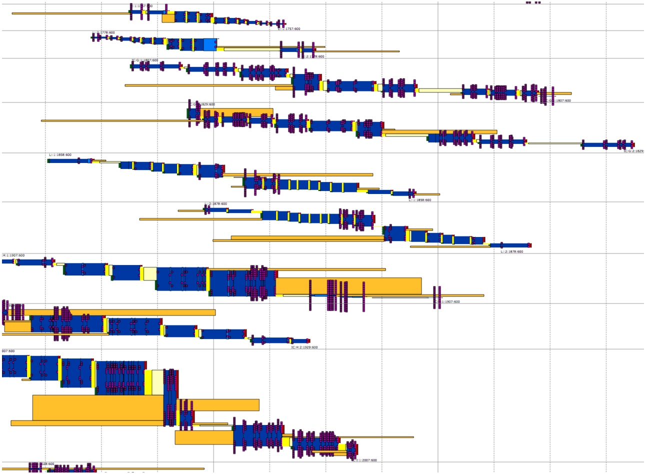
All Knock-On Costs for $N(N-1)$ Trains on Same Resource: Linearisation



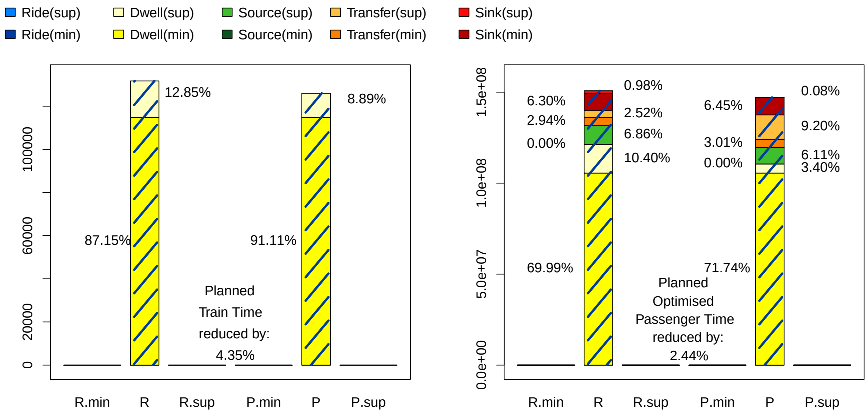
$$\forall R : \forall_{\substack{i,j \in I_R \\ i \neq j}} : \begin{cases} (s_{i,j,0}, ko_{i,j,0}) = (0, f_j \cdot \frac{a_j}{a_i(a_i+a_j)}) \\ (s_{i,j,1}, ko_{i,j,1}) = (T/15, f_j \cdot \frac{a_j e^{-a_i T/15}}{a_i(a_i+a_j)}) \\ (s_{i,j,2}, ko_{i,j,2}) = (T, f_j \cdot \frac{a_j e^{-a_i T}}{a_i(a_i+a_j)}). \end{cases} \tag{11}$$

$$\forall R : \forall_{\substack{i,j \in I_R \\ i \neq j}} : \begin{cases} pKO_{R,i,j} \geq ko_{i,j,0} + \frac{ko_{i,j,1} - ko_{i,j,0}}{s_{i,j,1} - s_{i,j,0}} \cdot (s_{i,j} - s_{i,j,0}) \\ pKO_{R,i,j} \geq ko_{i,j,1} + \frac{ko_{i,j,2} - ko_{i,j,1}}{s_{i,j,2} - s_{i,j,1}} \cdot (s_{i,j} - s_{i,j,1}) \end{cases} \tag{12}$$

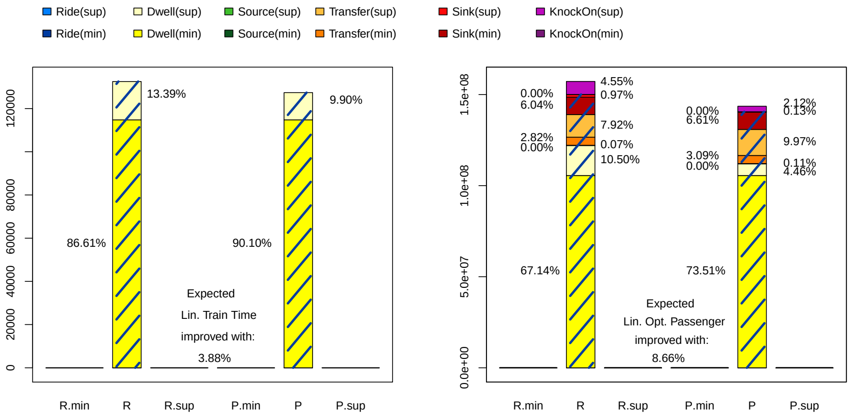
Results: Flow * Duration Rectangle Representation



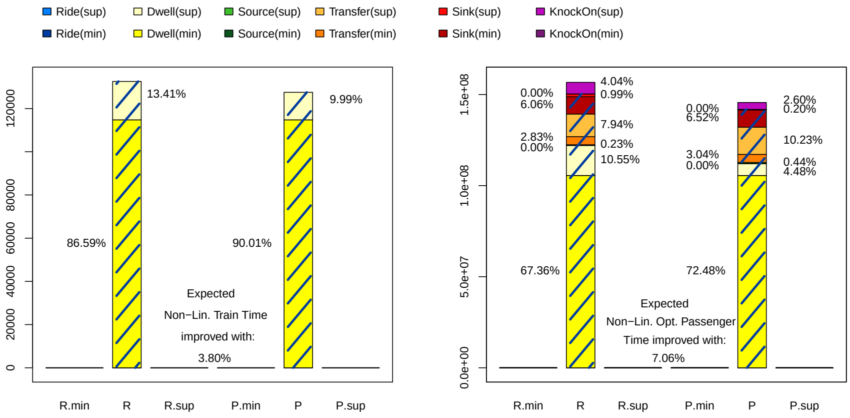
Planned Time



Expected Linear Time, as used in optimisation



Expected Non-Linear Time, as used in evaluation



Expected Linear Time, as used in optimisation

Table: Increasing primary delays, characterised by their average of *a*% of min. dwell & ride times. Graph size: 203 hourly trains, 5355 ride, 5152 dwell, 17553 major transfer, 31696 knock-on and 166 turn-around edges. Model size: 42609 supplement decision variables, 49415 integer decision variables, 41128 goal function terms for major flows and 58441 evaluation function terms for all flows.

a %	solver time min.	MILP gap %	major flows linearised ko-time reduction %	major flows linearised time reduction %	major flows non-linearised time reduction %	all flows linearised time reduction %	all flows non-linearised time reduction %
2	95	76.2	57	8.66	7.06	1.71	0.42
4	43	71.0	52	6.61	4.42	0.84	-1.41
6	75	61.3	63	7.65	5.73	2.07	0.13
8	66	61.3	59	5.83	3.86	0.40	-1.61
2	112	72.6	66	10.58	9.19	2.54	1.31

Conclusions

- defined and implemented remapping, reflowing, retiming & iterations
- reflowing: obtains local passenger numbers $\forall$ trains, $\forall$ locations
- retiming
  - defined all necessary constraints & found
    $\Rightarrow$ respects (ride, dwell, transfer, headway)-minimum times
  - added some our particular cycle set
    $\Rightarrow$ solves model fast
  - defined stochastic passenger time goal function
    - **derived & documented**
      Knock-On delay model for MILP timetable optmisation
      $\Rightarrow$ **ideal order and headway of trains**
      $\Rightarrow$ **ideal passenger robustness**
  - auto-generated first national timetable with full goal function = expected passenger time
  - reduction of passenger time with $\pm 7\%$, mind current assumptions:
    - primary delay = 2% of minimum-time, everywhere
    - zone-to-station-(overly?)-diffused passenger streams

Future Work

- further verification with new data
  - measured (place, train)-dependent delays i.o. averaged one
  - asymmetric station-OD?
- **add spreading measure for alternative OD-routes and evaluate effect**
- allow boundary timing conditions at frontiers/sub-zones
- output TPP problems to platformer
  - guarantee/increase chance on feasibility
    - add station capacity constraints to retiming
    - add constraints avoiding simultaneous arrival/departure of train pair that has to cross in station
  - adapt platformer so that it optimises for passengers i.o. maximising # trains platformed

Questions

- Your Questions?
  - www.LogicallyYours.com/Research/
  - sels.peter@gmail.com
- My Questions:
  - Is it best to use primary delays from the old timetable or to just assume them to be relative to minimum times?
  - If relative, what is the best (average(?)) percentage to assume for primary delays w.r.t minimum times?