From Data to Braking Curve

Analysis of Big Data Streams to obtain Braking Reliability Information for Train Protection system

Raphael Pfaff, Ingo Elsen, Bernd Schmidt

European Centre for Sustainable Mobility @ FH Aachen

FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

Problem setting

- Braking curves are fundamental to ATP systems
  - Exhibit random variation
- *A priori* braking curves
  - For limited configurations
  - So called γ-trains
  - By Monte Carlo simulation
- Freely configurable trains
  - So called λ-trains
  - Too many configurations for *a priori* determination
  - Handled by correction factors

Can we use a more current approach to determine braking curves?

Can we use a more current approach to determine braking curves?

Can we update braking curves for the observed braking performance?

FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

Can we use a more current approach to determine braking curves?

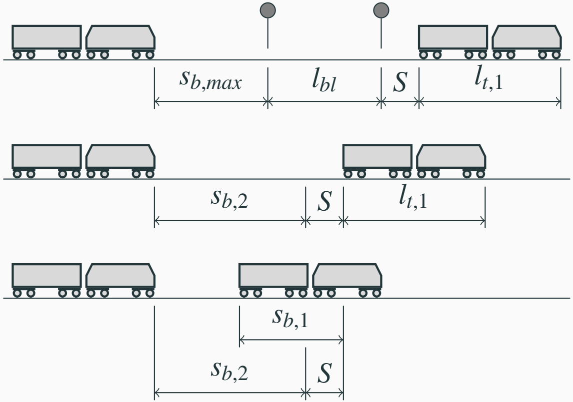
Can we update braking curves for the observed braking performance?

Can we perform *ad hoc* calculation of braking curves for freely configurable trains?

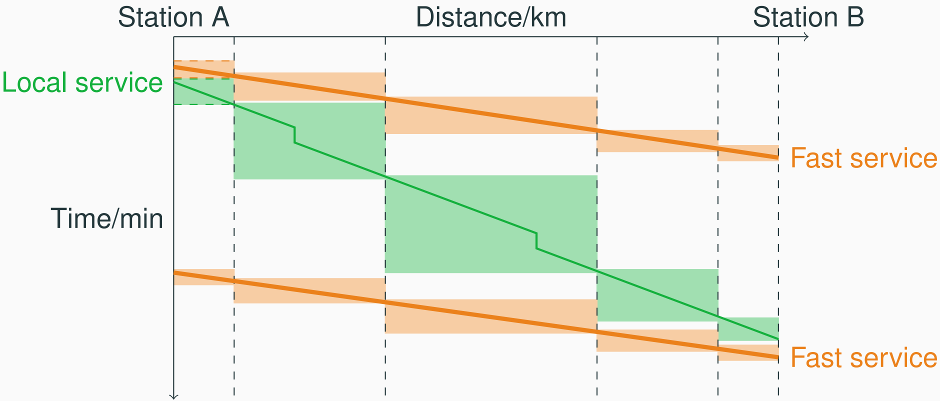
Introduction

Operation under ATP systems

- Requirements
  - safely locked
  - reserved
  - free from obstacles
- ATO system classification
  - spatially discrete
  - spatially continuous
  - mixed systems
- Headway
  - fixed spatial distance
  - absolute braking distance
  - relative braking distance



Interlocking and capacity improvements



Improvements in headway are likely to be consumed by bottlenecks in infrastructure or timetable.

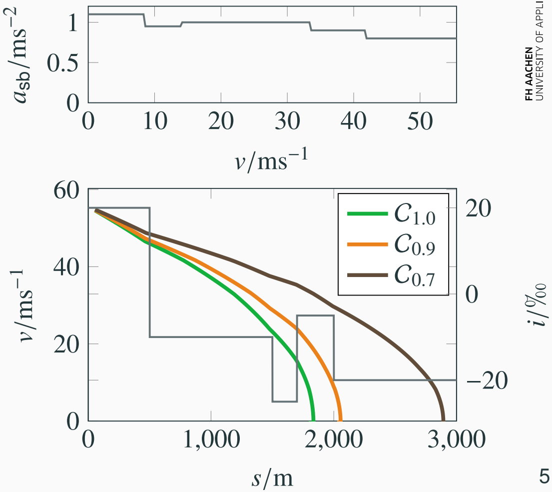
Train driver

- Drivers constantly monitor braking performance
  - Adhesion condition, by e.g.
    - Visual clues
    - Audible clues
    - Feeling
    - Experience on same/similar line
  - Vehicle condition, by e.g.
    - Exchange with colleagues
    - Experience with vehicle
- Drivers seek to optimise braking w.r.t.
  - Timetable, delays
  - Track and vehicle condition
  - Passengers

FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

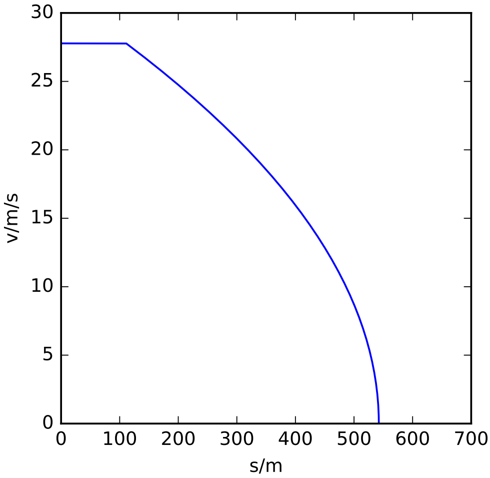
Train driver

- Drivers constantly monitor braking performance
  - Adhesion condition, by e.g.
    - Visual clues
    - Audible clues
    - Feeling
    - Experience on same/similar line
  - Vehicle condition, by e.g.
    - Exchange with colleagues
    - Experience with vehicle
- Drivers seek to optimise braking w.r.t.
  - Timetable, delays
  - Track and vehicle condition
  - Passengers



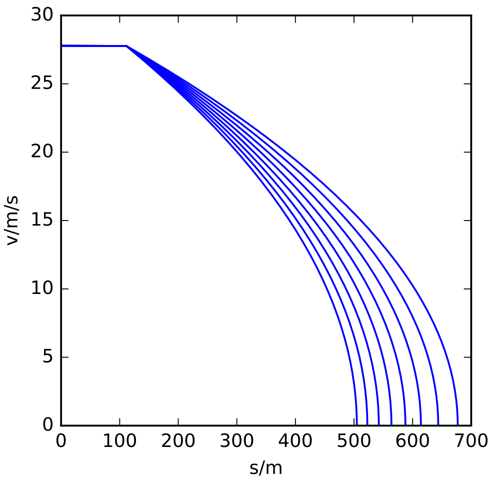
Braking curves

- To supervise train velocity, ATP systems predict the future braking capability of the train



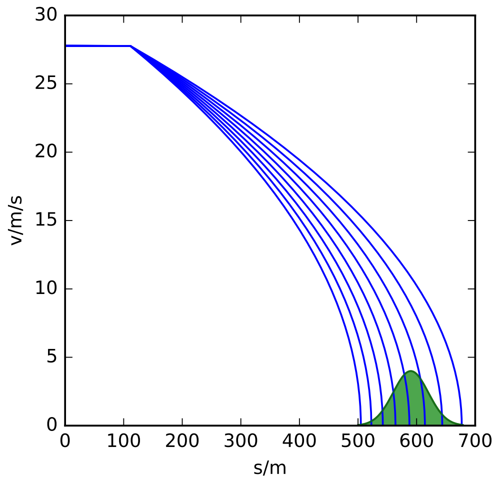
Braking curves

- To supervise train velocity, ATP systems predict the future braking capability of the train
- However, there is not *the* braking capability



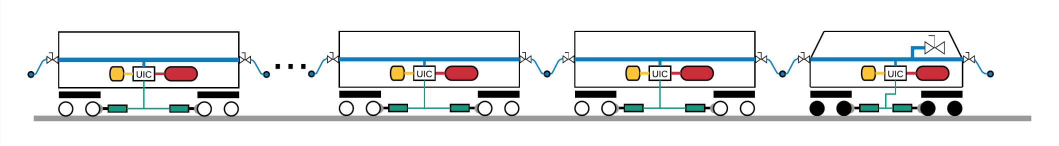
Braking curves

- To supervise train velocity, ATP systems predict the future braking capability of the train
- However, there is not *the* braking capability
- Braking curves exhibit a random behaviour



Random braking curve behaviour

- Brake pipe: propagation velocity, flow resistances, train length
- Distributor valve: Filling time, brake cylinder pressure
- Braking force generation: efficiency, brake radius (for disc brakes), pad/block friction coefficient
- Wheel/rail contact: rail surface, contaminants, slip, ...
- Also discrete failure events need to be considered



Big Data and Data Streams

Big Data

- Dominating railway applications:
  - Predictive maintenance
  - Improvement of operations
- Classification according to different sets of V, e.g.
  - Volume: quantity of stored and generated data
  - Variety: type and nature of the data
  - Velocity: speed at which the data is generated
  - Veracity: quality of the captured data

Camelia.boban (CC BY-SA 3.0)

Expected data volume

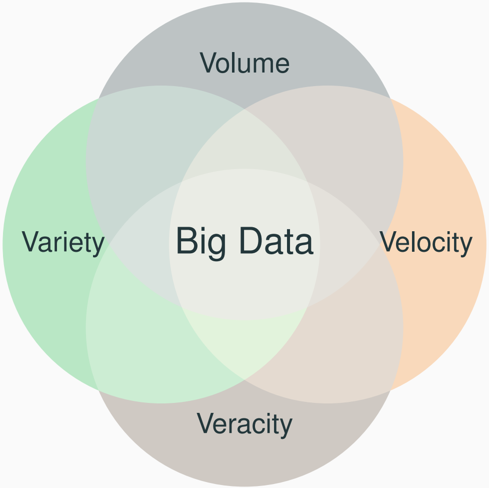
- Using DB open data:
  - 2.6 million monthly freight trains
- Assumptions:
  - 10 braking processes per train
  - 20 wagons per train
- Data acquisition:
  - 90 s braking duration
  - 10 Hz sampling rate
  - 20 bytes per sample

Expected data volume (for DB only) 9.4 TB per month

Month	Freight trains operated
February	2,641,295
March	2,712,662
April	2,734,730
May	2,497,157
June	2,719,753
July	2,576,785
August	2,472,119
September	2,660,830
October	2,513,284
November	2,641,139

Expected data velocity, variety and veracity

- Velocity: 7.2 MB/s
  - Expected data volume (for DB cargo only) 9.4 TB per month
  - Mostly night operation: 360 h/month
- Variety: medium
  - Highly correlated with track, weather and maintenance data
  - Needs to be collected as much as possible
- Veracity: medium to low
  - Low sensor cost required
  - Workshops not capable of checking or calibrating

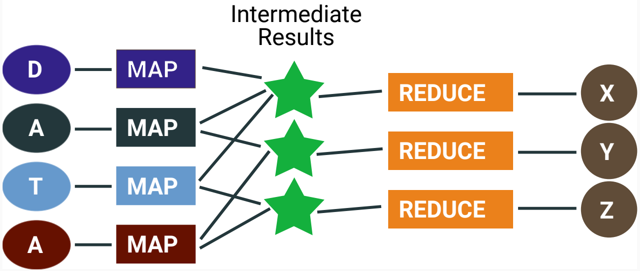


Processing slow and fast data

FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

Data processing architecture

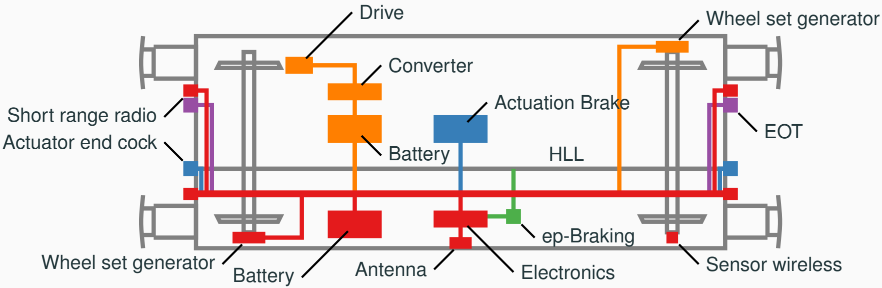
- Requirements similar to real-time arrival prediction
  - Sub-second latency
  - Full pipeline few seconds latency
- Architectural elements:
  - Streaming layer: Lambda architecture, e.g. Apache Storm
  - Processing Layer: Hadoop-Tools, e.g. Hive, or MapReduce



Vehicle implementation

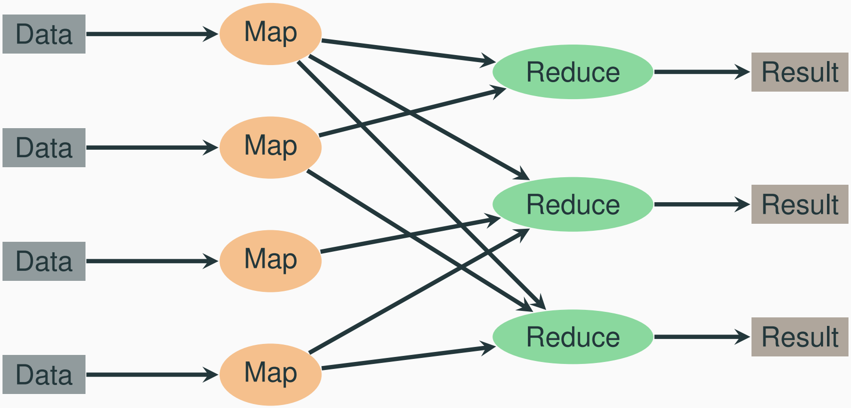
Wagon 4.0 as enabler

- Wagon 4.0 concept pillars:
  - Power supply
  - Sensors
  - Actuators
  - Algorithms
  - WagonOS operating system



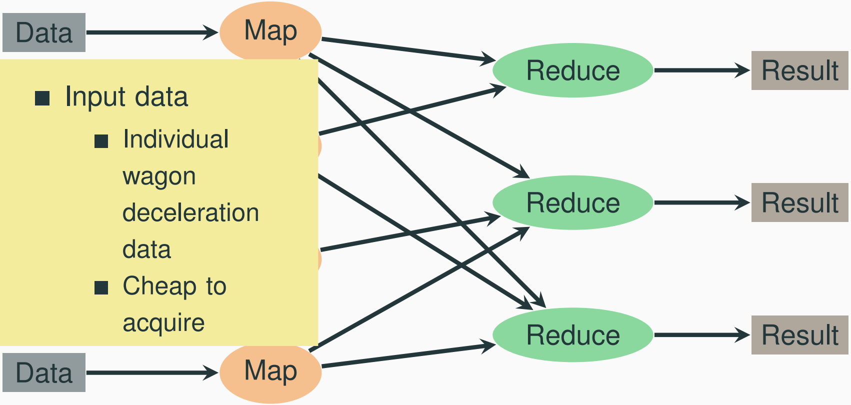
FH AACHEN UNIVERSITY OF APPLIED SCIENCES

Proposed algorithm



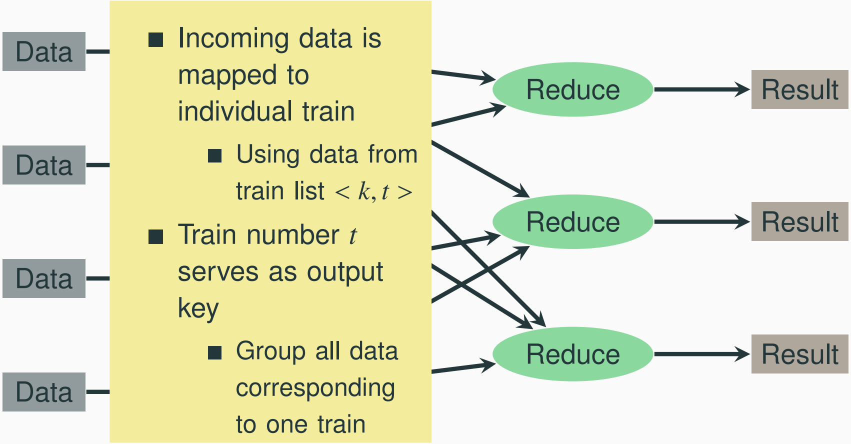
FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

Proposed algorithm

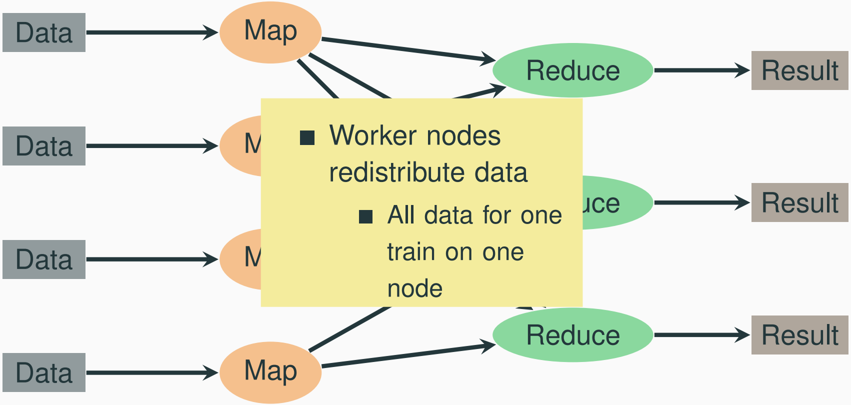


- Input data
  - Individual wagon deceleration data
  - Cheap to acquire

Proposed algorithm



Proposed algorithm



- Worker nodes redistribute data
  - All data for one train on one node

Proposed algorithm

Data

Data

Data

Data

Observed braking force

$$F_{B,obs} = \frac{a}{m_i} \tag{1}$$

Bayesian update of prior information

$$\sigma_{est}(k) = \alpha \sqrt{\frac{\sigma^2 \sigma_{est}(k-1)^2}{\sigma^2 + \sigma_{est}(k)^2}} \tag{2}$$

$$F_{B,est} = \frac{F_B \sigma^2 + F_{B,obs} * \sigma_{est}(k)^2}{\sigma^2 + \sigma_{est}(k)^2} \tag{3}$$

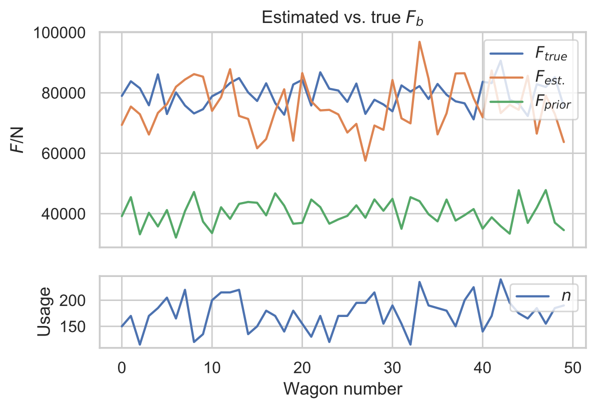
Result

Result

Result

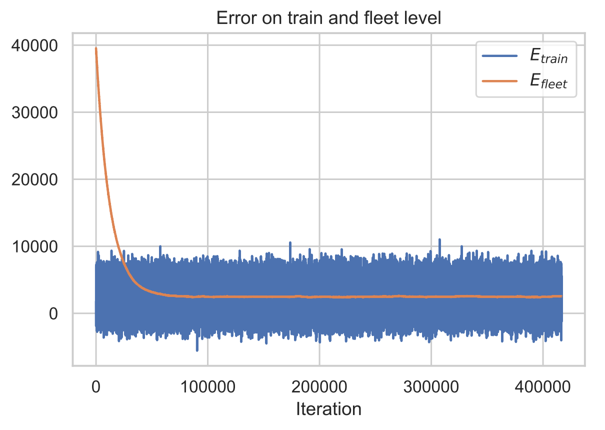
Numerical example

- Rolling stock
  - $N = 5 \cdot 10^4$ wagons
  - $n = 20$ wagons/train
- Operation
  - $M = 83.3 \cdot 10^4$ trains/day
  - Randomly assigned from wagon pool
- Simulation
  - Initialised conservatively at 50% performance
  - Additive white noise on sensor signal
  - Implemented in Python



Numerical example

- Rolling stock
  - $N = 5 \cdot 10^4$ wagons
  - $n = 20$ wagons/train
- Operation
  - $M = 83.3 \cdot 10^4$ trains/day
  - Randomly assigned from wagon pool
- Simulation
  - Initialised conservatively at 50% performance
  - Additive white noise on sensor signal
  - Implemented in Python



FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

Conclusion and further work

- Conclusion:
  - *Ad hoc* estimation of braking curve appears feasible
  - Observing true deceleration may yield improvements over current procedures
  - Proposed algorithm employs the well known MapReduce framework for scalability
- Further work:
  - Apply to real world data
  - Include rich data sources, such as weather, maintenance performed etc.
  - Aiming to resemble train driver behaviour as closely as possible

SPONSORED BY THE

Federal Ministry of Education and Research

FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

Let’s put awesome back into railways!

Prof. Dr. Raphael Pfaff · pfaff@fh-aachen.de · +49-151-70052454 · www.r10f.com