EXAMINING CELLULAR ACCESS SYSTEMS ON TRAINS: MEASUREMENTS AND CHANGE DETECTION

JOHAN GARCIA, ***STEFAN ALFREDSSON***, ANNA BRUNSTROM
DEPARTMENT OF COMPUTER SCIENCE
KARLSTAD UNIVERSITY, SWEDEN

COMPUTER SCIENCE
DATAVETENSKAP

OUTLINE

- Problem formulation
- Train cellular access setup
- Measurement collection
- Overview of velocity, throughput, RTT distributions

- Changepoint detection
- Sigmoid fitting and changepoint coefficient functions
- Examples of detected changepoints in dataset
- Conclusions and future work

PROBLEM FORMULATION

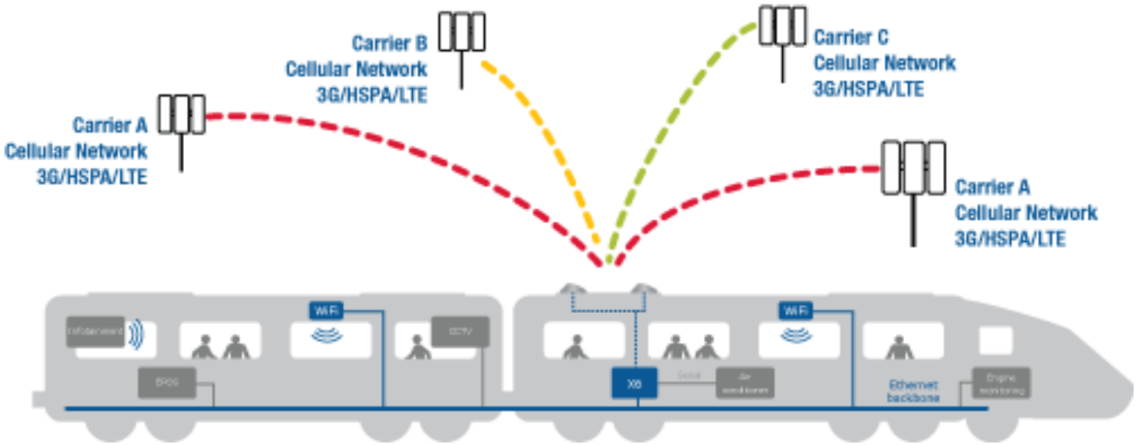
- How does mobile broadband networks perform on high-speed railway trains using aggregation routers?
- How can anomalies be detected?

- Extensive dataset from railway train aggregation routers
- Changepoint detection

TRAIN CELLULAR ACCESS SETUP

Each train has an aggregation router, providing WiFi access to passengers.

The backhaul consists of four aggegated LTE modems, connected to two operators.



MEASUREMENT COLLECTION

- Trains traveling Stockholm-Göteborg, Stockholm-Malmö, Stockholm-Karlstad
- In total over 7000 journeys
- Data collected passivly every 5 seconds
    - Number of active users
    - Link throughput
    - ICMP RTT
    - GPS
    - Network characteristics

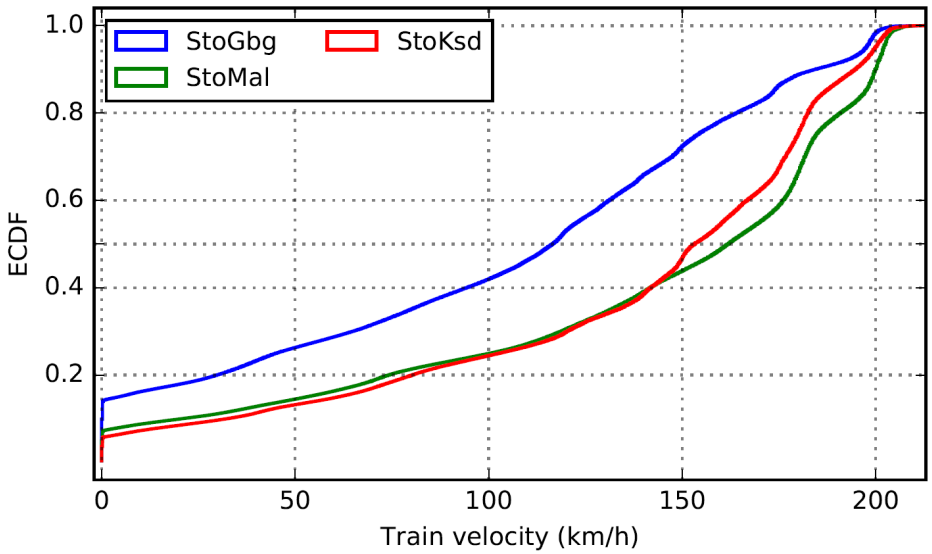
Line name	Track Length	Nr of journeys	Nr router ids	Nr of cellids	Avg velocity
StoGbg	485 km	2380	54	5589	103 km/h
StoMal	615 km	3900	36	5048	137 km/h
StoKsd	325 km	1458	52	2765	135 km/h
Overall		7738	97	11644	125 km/h

CELLULAR LINK TECHNOLOGY DISTRIBUTION

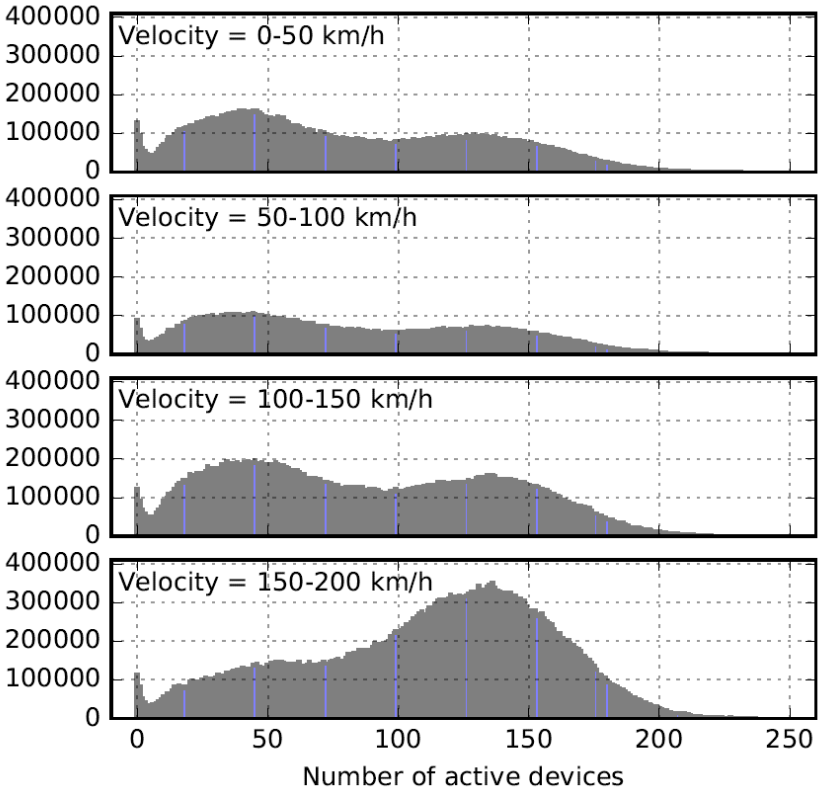
Link tech. (%)	StoGbg		StoMal		StoKsd	
	Op1	Op2	Op1	Op2	Op1	Op2
LTE	99.979	72.333	99.977	96.667	99.814	90.814
HSPA+	0.000	18.506	0.000	2.462	0.001	8.396
DC-HSPA+	0.018	3.186	0.009	0.336	0.128	0.526
HSDPA	0.001	2.848	0.003	0.323	0.006	0.076
HSPA	0.001	2.295	0.003	0.149	0.018	0.117
UMTS	0.000	0.518	0.000	0.044	0.002	0.033
Nr obs.	15.79MR	15.61M	24.65M	24.86M	4.91M	4.97M

- In total over 90 million measurements
- Mainly LTE for Op 1, Op 2 also uses other technologies

TRAIN VELOCITY DISTRIBUTIONS

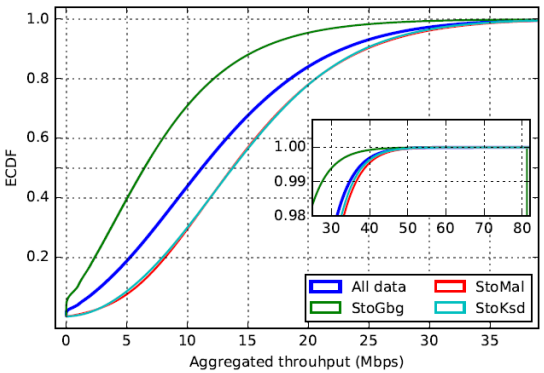


Velocity per train line

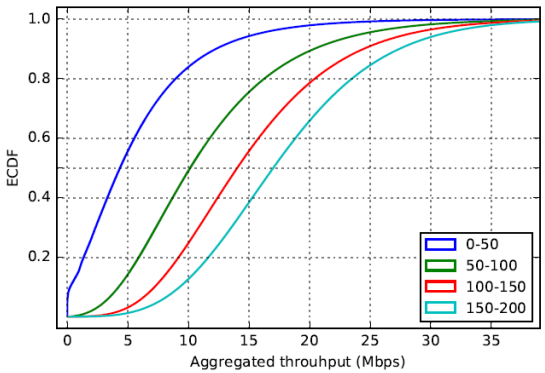


Active devices per velocity band

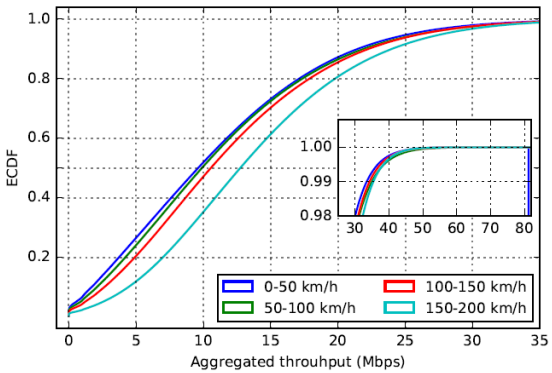
AGGREGATED DOWNLINK THROUGHPUT



(a) per train line



(b) per interval of active devices



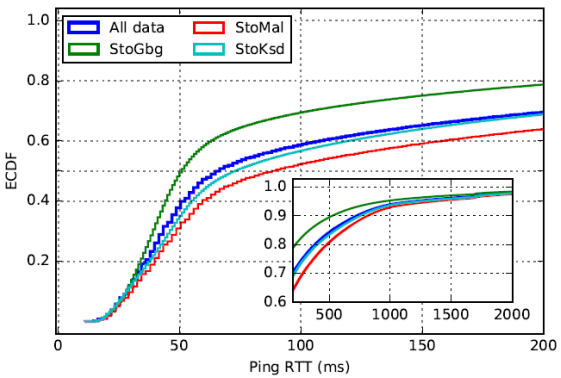
(c) per velocity intervals

A: The Gothenburg run has less throughput vs Malmö and Karlstad runs

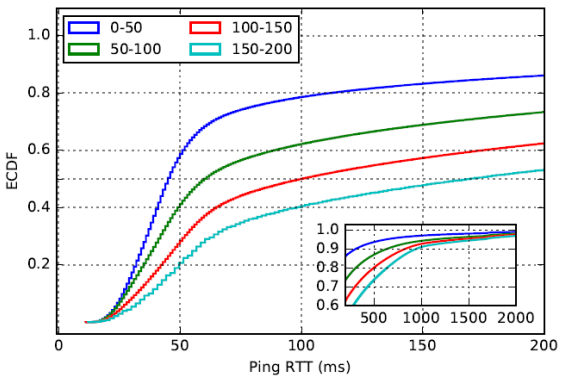
B: Aggregate throughput increases with number of devices

C: Large coupling between aggregated throughput and velocity

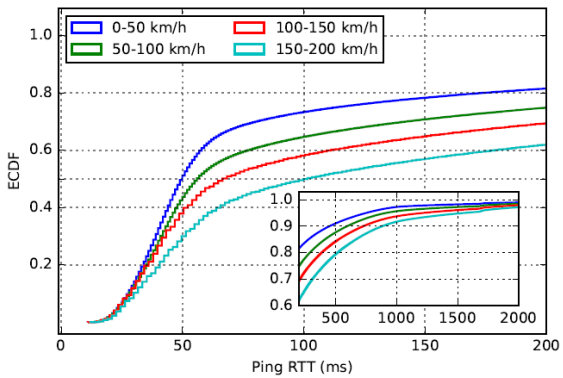
PING RTT



(a) per train line



(b) per interval of active devices



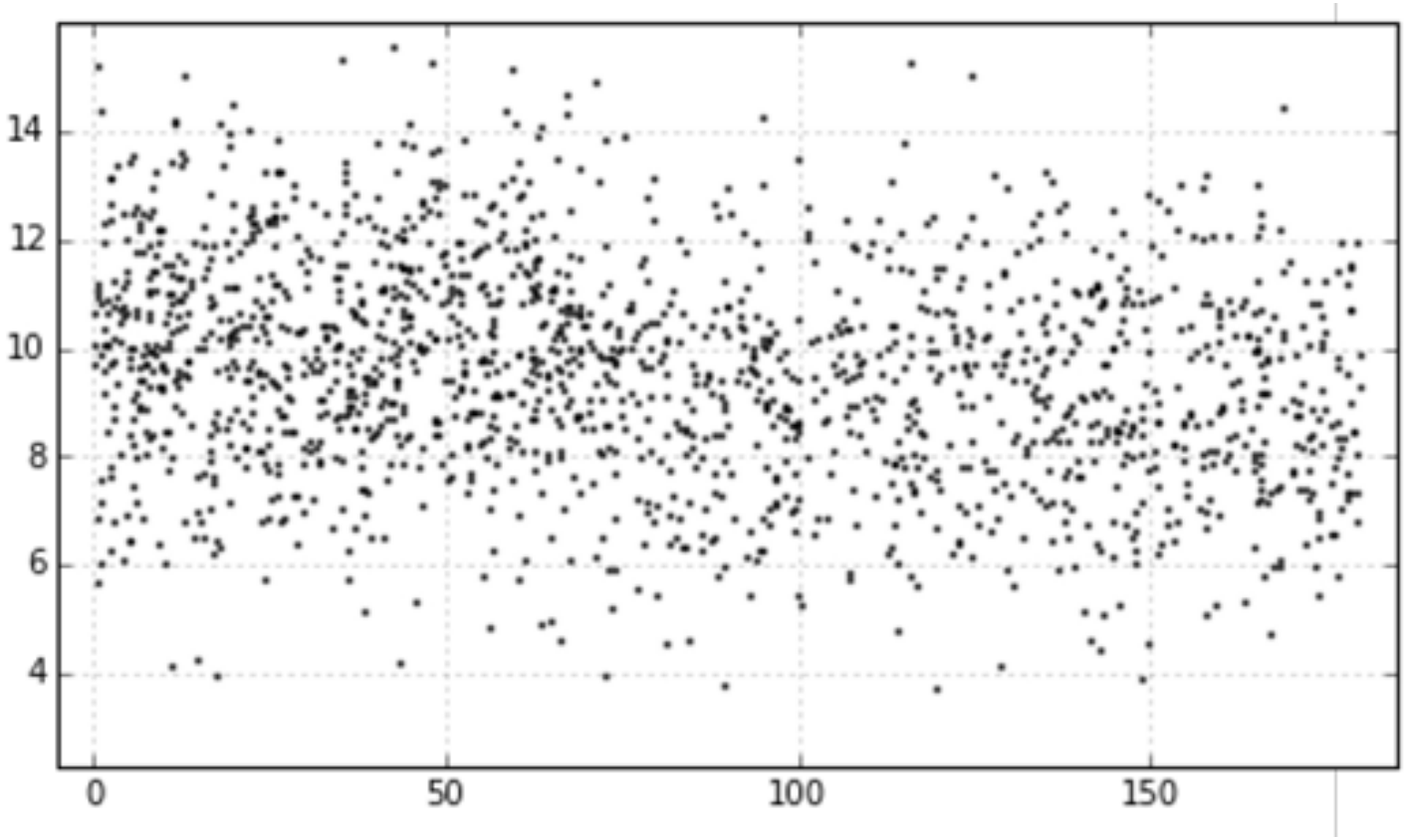
(c) per velocity intervals

Overall RTT per train line, per active device and velocity interval.

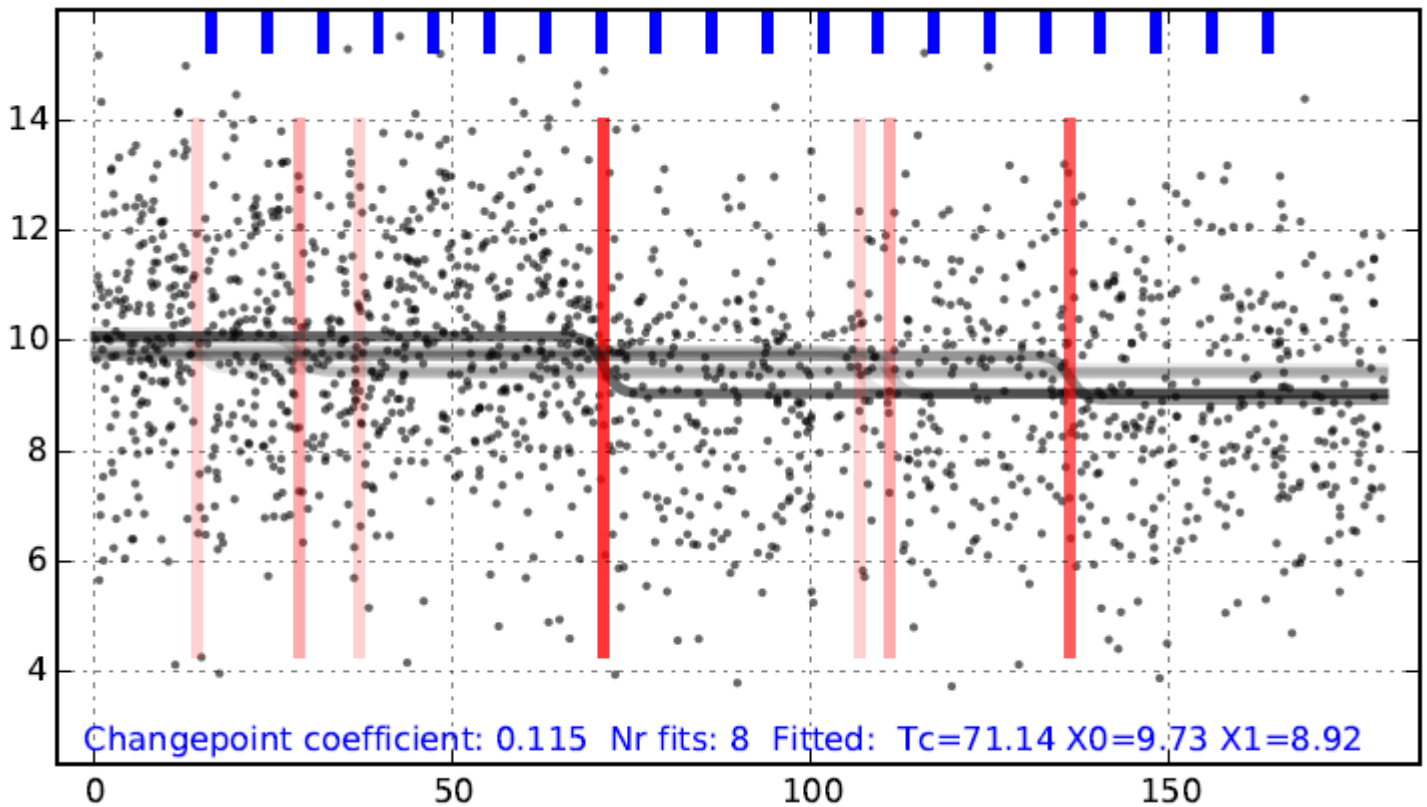
Consistent with throughput/load, indicating excessive buffering/bufferbloat

Changepoint Detection

CHANGEPOINT DETECTION EXAMPLE



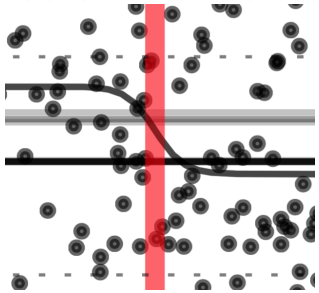
SIGMOID FITTING CHANGEPOINT DETECTION



SIGMOID AND CHANGEPOINT COEFFICIENT FUNCTIONS

The sigmoid function:

- S-shaped
- Two stationary means with a changeover period
- Fitting via non-linear least squares Levenberg-Marquardt regression
- 20 runs with different Tc seed values



$$Y(T) = \frac{Y_1 - Y_0}{1 + e^{-S*(T-T_C)}} + Y_0$$

Changepoint coefficient

- Represents the relative amount of change at Tc
- Maximum C is used for changepoint candidate

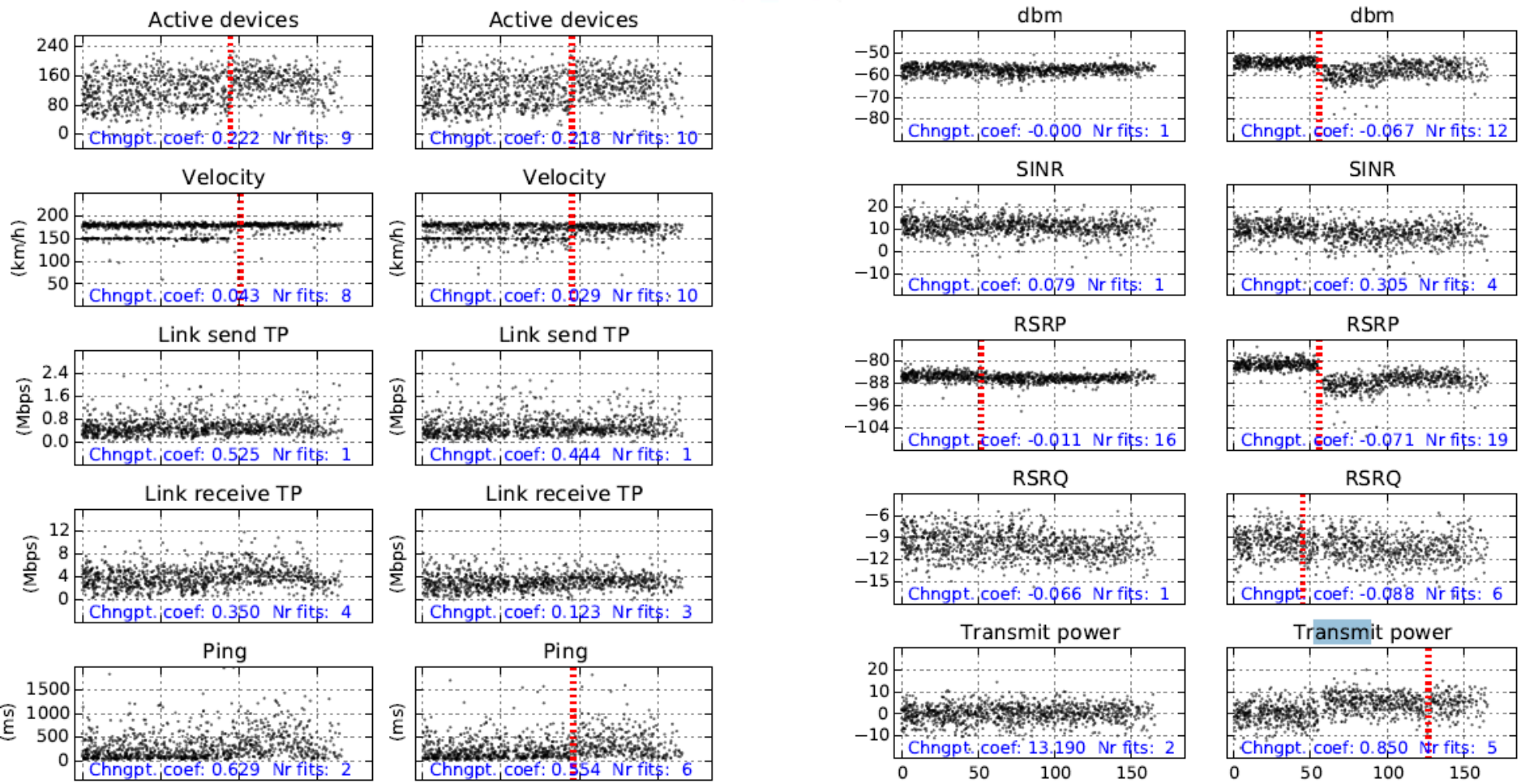
$$C(T_c) = \frac{|Y_1 - Y_0|}{\min(Y_0, Y_1)}$$

APPLYING THE CHANGE POINT DETECTION

With a subset of 200 cells crossed by the train on the Stockholm-Karlstad stretch

1. Compute aggregated averages per cellid-trainpassing event
2. Use aggregates to compute 20 LM sigmoid fittings
3. Find time locations of highest changepoint coefficient
4. Threshold on the number of LM fittings with the highest changepoint cofficient (5 out of 20)

EXAMPLE OF DETECTED CHANGEPOINTS



COMPUTER SCIENCE
DATAVETENSKAP

CONCLUSIONS AND FUTURE WORK

- Analysis of dataset from operational Swedish railway system used by passengers
    - Shows strong correlation between number of users and aggregated throughput
    - Ping RTT highly dependent on users and throughput, indicating buffer bloat issues
    - Changepoint detection to find possible anomalies
- Future work:
    - Refining the changepoint detection approach
        - Outlier rejection, Box-Cox transformation, data scaling
    - Complement passive measurements with additional active measurements

Questions?