EXAMINING CELLULAR ACCESS SYSTEMS ON TRAINS: MEASUREMENTS AND CHANGE DETECTION

JOHAN GARCIA, **STEFAN ALFREDSSON**, ANNA BRUNSTROM
DEPARTMENT OF COMPUTER SCIENCE
KARLSTAD UNIVERSITY, SWEDEN
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
Each train has an aggregation router, providing WiFi access to passengers. The backhaul consists of four aggregated LTE modems, connected to two operators.
MEASUREMENT COLLECTION

- In total over 7000 journeys
- Data collected passively every 5 seconds
  - Number of active users
  - Link throughput
  - ICMP RTT
  - GPS
  - Network characteristics

<table>
<thead>
<tr>
<th>Line name</th>
<th>Track Length</th>
<th>Nr of journeys</th>
<th>Nr router ids</th>
<th>Nr of cell ids</th>
<th>Avg velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>StoGbg</td>
<td>485 km</td>
<td>2380</td>
<td>54</td>
<td>5589</td>
<td>103 km/h</td>
</tr>
<tr>
<td>StoMal</td>
<td>615 km</td>
<td>3900</td>
<td>36</td>
<td>5048</td>
<td>137 km/h</td>
</tr>
<tr>
<td>StoKsd</td>
<td>325 km</td>
<td>1458</td>
<td>52</td>
<td>2765</td>
<td>135 km/h</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>7738</td>
<td>97</td>
<td>11644</td>
<td>125 km/h</td>
</tr>
</tbody>
</table>
CELLULAR LINK TECHNOLOGY DISTRIBUTION

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

- 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: 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

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

Changepoint coefficient: 0.115  Nr fits: 8  Fitted: Tc=71.14 X0=9.73 X1=8.92
SIGMOID AND CHANGEPPOINT 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

Changepoint coefficient
• Represents the relative amount of change at Tc
• Maximum C is used for changepoint candidate

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

\[ C(T_c) = \frac{|Y_1 - Y_0|}{\min(Y_0, Y_1)} \]
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 coefficient (5 out of 20)
EXAMPLE OF DETECTED CHANGEPOINTS

Active devices

Velocity

Link send TP

Link receive TP

Ping

Active devices

Velocity

Link send TP

Link receive TP

Ping

dbm

SINR

RSRP

RSRQ

Transmit power

Chngpt. coef: 0.222  Nr fits: 9

Chngpt. coef: 0.218  Nr fits: 10

Chngpt. coef: 0.043  Nr fits: 8

Chngpt. coef: 0.029  Nr fits: 10

Chngpt. coef: 0.529  Nr fits: 1

Chngpt. coef: 0.444  Nr fits: 1

Chngpt. coef: 0.350  Nr fits: 4

Chngpt. coef: 0.123  Nr fits: 3

Chngpt. coef: 0.629  Nr fits: 2

Chngpt. coef: 0.554  Nr fits: 6

Chngpt. coef: 0.000  Nr fits: 1

Chngpt. coef: 0.067  Nr fits: 12

Chngpt. coef: 0.079  Nr fits: 1

Chngpt. coef: 0.305  Nr fits: 4

Chngpt. coef: 0.911  Nr fits: 16

Chngpt. coef: 0.671  Nr fits: 19

Chngpt. coef: 0.066  Nr fits: 1

Chngpt. coef: 0.088  Nr fits: 6

Chngpt. coef: 13.190  Nr fits: 2

Chngpt. coef: 0.850  Nr fits: 5
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?