Examining Cellular Access Systems on Trains: Measurements and Change Detection

Johan Garcia, Stefan Alfredsson, Anna Brunstrom
Department of Mathematics and Computer Science, Karlstad University, Sweden
Email: {johan.garcia, stefan.alfredsson, anna.brunstrom}@kau.se

***Abstract*—Access to reliable high-quality communication services on trains is important for today's mobile users. Train-mounted aggregation routers that provide WiFi access to train passengers and bundle external communication over multiple cellular modems/links is an efficient way of providing such services. Still, the characteristics of such systems have received limited attention in the literature. In this paper we examine the communication characteristics of such systems based on a large data set gathered over six months from an operational Swedish railway system. We characterize the conditions in terms of usage load, train velocity profiles, and observed throughput and delay as well as the relation between these parameters. Furthermore, we examine the data from an anomaly detection perspective. Based on a changepoint detection method, we examine how the collected metrics varies over the six months. Being able to detect shifts in the metrics over time can help detect anomalous changes in the hardware or environment, and also further helps explain the factors affecting the observed behaviors.**

I. Introduction

In todays society most personal communication devices are multi-radio and capable of both cellular 3G/4G, as well as WiFi, connectivity. Being portable, these devices naturally follow along wherever a person goes. A common transportation is railway travel. This mode of transport brings interesting challenges for mobile communication. At high vehicular speeds there are frequent handovers, and rapidly changing signal quality as the signal fades by the distance to the cell towers, and also experiences signal multipathing by objects in the environment. This is further aggravated by the number of persons on a train, which can each carry a number of devices, which increases the load on the network infrastructure.

Instead of having each device communicate with the cellular infrastructure, they can instead be directed to a train-local access network, or aggregation router. This router provides clients with WiFi access, and traffic is then transported in aggregate via roof-top antennas for better coverage. Such aggregation routers can also utilize multi-operator connectivity transparently to the client, which can enable communication resiliency in case of blind spots for one or the other operator. An illustration of such a system is provided in Figure 1.

This paper analyzes train communication based on real-life measurements captured at train aggregation routers. It provides two main contributions. First we provide an overall characterization of train journey details, communications usage, radio technology and handover distributions along with observed system throughput and delay characteristics. As the system can be bounded by limited aggregate demand, there is a visible dependency between the number of active users and throughput and delay. Secondly, we highlight longitudinal effects in the data, using a sigmoid-fitting changepoint detection approach. The approach is shown to be able to automatically detect relevant changes in metric behavior over time. In the following, we describe the related work in section II, followed by a description of the collection and characteristics of the dataset (sections III and IV) which form the basis for the changepoint analysis in section V. Finally the conclusions are presented in section VI.

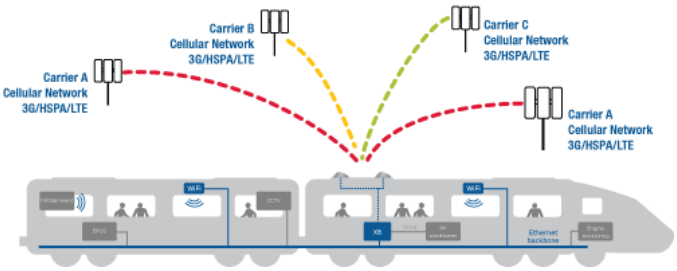


Figure 1: Train communication system based on a cellular router structure [courtesy of Icomera AB]

II. Related Work

Our work involves measurements in high-velocity railway communication scenarios using LTE and its characteristics. Other studies in this environment include for example [1], [4], [7], [10], [11], [12]. These studies are primarily concerned with radio level aspects, or using LTE in general, for enabling connectivity on trains. Access can be provided in a number of ways to users, for example via direct connection to the users, via local base stations, or proxy access terminals. A survey of different solutions is provided in Masson et al. [9], where the authors remark that train access terminals "represents the best technical solution to optimize performance and throughput". A train access terminal then provides the users with WiFi access, routed through an aggregation gateway that itself uses mobile broadband or satellite connections. For this paper, we perform studies of such an access terminal system consisting of an aggregation gateway with WiFi on the LAN side and LTE on the WAN side. In further related work, Mueller et al. [11] compare the performance of using a train access terminal versus users having a direct connection, by means of simulation studies. The results show that access terminals

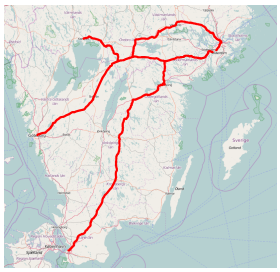


Figure 2: Train routes

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

Table I: Train route characteristics

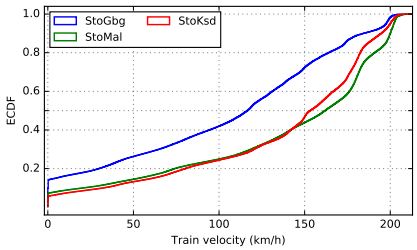


Figure 3: Route velocity distribution

provide the best performance, but individual connectivity is also performing well.

Handling of handovers is an issue with train-based communication. By using dual radios, the handover time can be improved where base station coverage is overlapping, as shown in an analytical study by Lin et al. [7].

The radio link performance at high velocities is studied by Merz et al. [10]. They use two traces collected from operational trains, with several hours of trace time. The most important factor for reliable operation is shown to be the signal to noise ratio. In [6], Li et al. identify problems on high-speed trains, revealing that disconnections, dropped packets and RTT spikes cause severe problems for TCP. Lutu et al. [8] perform measurements on Norwegian trains. Based on measurements of download performance and packet loss, they provide a mapping of operators and their coverage. These measurements are using individual connections of user equipment to the operator infrastructure. In contrast, our work studies the aggregate traffic of all users connected via the access terminal, and then aggregated over a number of uplink 3G/4G connections to multiple operators.

III. Data Set Overview

A. Data set collection

The underlying data for this study is trace measurements collected from live train journeys during an extended time period (6 months) in 2016. They are from three different train routes, between the Swedish cities Stockholm-Göteborg, Stockholm-Malmö and Stockholm-Karlstad. Figure 2 shows a map with the outline of the routes. From these routes, data from over 7000 train journeys are used, from 97 unique routers, each corresponding to an individual train set. Table I shows the details of the routes. Among the data collected by the modems is the cellid. As outlined in the table, the average number of unique cellids per kilometer of track varies between the different stretches. For example Stockholm-Malmö has 8.2 cellids per km, while Stockholm-Göteborg has 11.5. For performance, it is however more important how the cells are geographically clustered, and their distance to the track, rather than the average density.

Data collection occurs every five seconds, and is done by the on-board router. The data collection includes for example how many active devices that are connected, the GPS position of the train and the current velocity. Data concerning radio-related metrics, aggregated throughput and active latency measurements using ping are also included and plays a key role in the analysis. The onboard router on each train has four mobile broadband modems (Sierra Wireless MC7710) with LTE capability. Two cellular operators (Op 1 and Op 2) are contracted for the connectivity with two modems connected to each. Traffic scheduling over the modems and operators is handled by the on-board router, using a proprietary scheme based on monitoring of individual link conditions and characteristics.

B. Train and usage characteristics

Train velocity is a factor that may influence the ability to provide cellular communication as higher speed means more handovers per unit of time. Figure 3 shows the distribution of train speeds. It can be observed that the ordering of the medians are consistent with Table I, and that the medians are larger than the average as would be expected due to the train being stationary during stops.

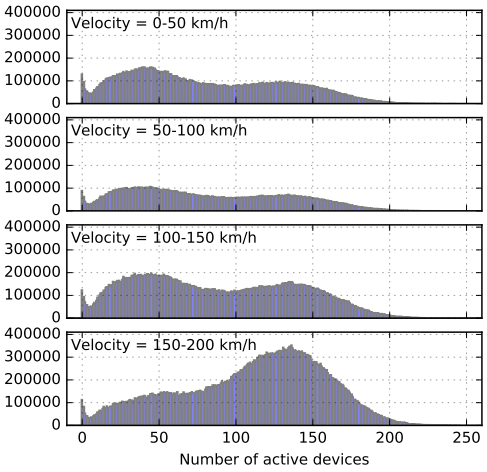


Figure 4: Histogram of active devices per velocity interval

A factor of interest is also the load to the system in terms of number of active communication devices. To gain access to the cellular connectivity, the user of each device has to accept the

service agreement, and the device is by that considered active and has the ability to use the provided communication services. The service policy in use at the time of data collection allowed each user 200MiB of download data per journey without capping the rate for the user. After 200MiB, the rate is capped at 0.5Mbps. The number of active devices can be expected to have some coupling to the train velocity as during the initial, lower speed, phase of the journey not all of the communication users will have authorized themselves to the train network. The distribution of active devices in relation to four intervals of train velocity is shown in Figure 4. It is clearly visible that the highest velocity interval has the majority of the measurements, which is consistent with Figure 3, and also that the highest velocity interval has a larger fraction of the measurements for high number of active devices as compared to the other velocity intervals. However, for all velocity intervals, a sizable fraction of measurements are collected with a load of 150 or more active devices.

IV. Measured overall characteristics

A. Cellular aspects

The modems report the cellular technology which they are connected to at each measurement collection. This data can be seen as a snapshot of the current state of technology deployment in the geographical regions. The distribution of cellular technologies per line and operator is shown in Table II. Out of the over 90 million link measurements 93.8% were using LTE. On a per operator basis, Op 1 obtained 99.9% LTE and Op 2 87.7% LTE. As the router selects which links to use based on the link conditions, the link aggregation algorithm might chose not to use a slower technology although connectivity exists.

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

Table II: Observations of link technology per operator

The handover frequency is another aspect of interest. As there are four modems in each train, the maximum number of changed cellids between consecutive observations that can occur is 4, when handovers has been performed at least once for all four modems. Table III shows the distribution of handovers per line.

Modems with cellid changes:	0	1	2	3	4
Stockholm-Göteborg	59%	24%	13%	3.6%	0.77%
Stockhlm-Malmö	55%	23%	16%	4.2%	1.5%
Stockholm-Karlstad	56%	23%	15%	4.4%	1.4%

Table III: Cellid changes since previous observation

B. Aggregated throughput and delay aspects

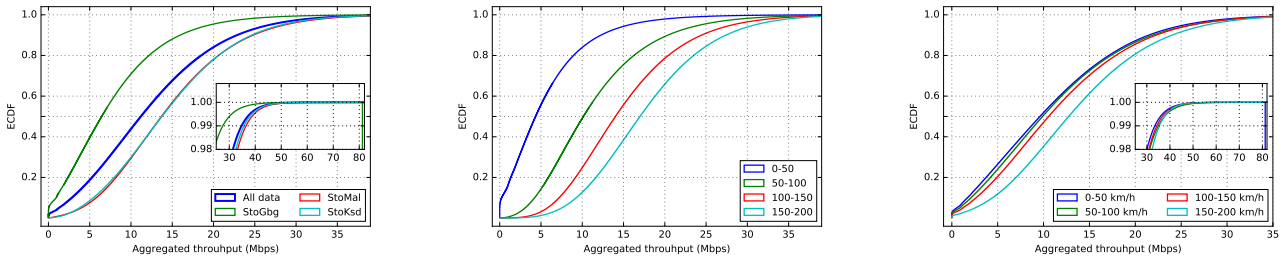
Now considering the achieved aggregate throughput, this will be bounded by the lesser of cellular radio constraints and aggregate demand from the users. The observed throughput can be bounded by the aggregate demand since these measurements are passive. Figure 5 shows the ECDFs of the aggregated throughput in relation to other metrics. The relationship between number of devices and throughput is shown in Figure 5b. In this figure it is apparent that there is a very large coupling between the number of devices and the aggregate throughput. The coupling between velocity interval and aggregate throughput is shown in Figure 5c. Some differences are observed for the different velocity intervals. However, this variation is to a large extent a reflection of the variation in number of active devices per velocity interval as shown in Figure 4 rather than an effect of the variation in the speed itself.

Another metric of interest is the ping round trip time (RTT), which is measured on a per link basis from the aggregation router to specific internet hosts. The overall observed ping RTT per train line is shown in Figure 6a. As can be seen in the figure, the ping times vary greatly from just over 10 ms to well over 2000 ms for all lines. We can further see that the line with the lowest aggregated throughput, StoGbg, also has the lowest ping times. As the throughput is connected to the load, this suggests that the ping times are also strongly correlated to the load in the system. This is confirmed by Figure 6b which shows the impact of offered load on the observed round trip times, displaying the ping times per active device quantile. For the quantile with the highest load, 50% of the ping times are over 150 ms, while less than 20% of the ping times are above 150 ms for the quantile with the lowest load. Similar as for the aggregate throughput, differences in load is very likely the main underlying cause for the difference in ping times seen for the different velocity intervals in Figure 6c.

The ping times in Figure 6, and their relation to load, would indicate that the amount of buffers in the cellular infrastructure can be substantial, and the connections suffer from considerable bufferbloat. As the amount of concurrent downlink traffic increases, the buffers in the cellular network fills up causing high delays for the ping traffic. This is consistent with previous studies of cellular networks that have also observed substantial bufferbloat in the presence of concurrent traffic [5], [2]. For all train lines, it is clear from Figure 6 that running interactive applications, like VoIP, could be challenging during parts of many train journeys.

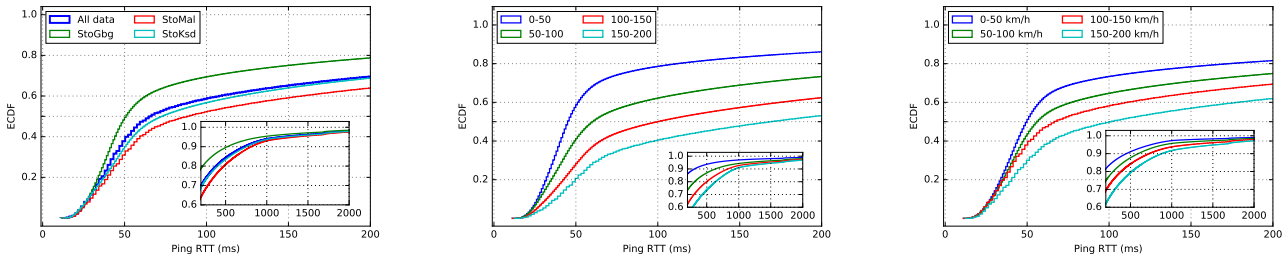
V. Changepoint Detection

In the previous section the characteristics of several metrics were provided, providing an overall intuition for the range of values present, and some of their interrelationships. In this section we instead focus on examining data from a longitudinal perspective. As the data was collected over a six month period, data stationarity aspects can be explored. Being able to detect shifts over time in the collected metrics can be very useful, both to detect anomalous changes in hardware or environment,



(a) per train line (b) per interval of active devices (c) per velocity intervals

Figure 5: Aggregated downlink throughput characteristics



(a) per train line (b) per interval of active devices (c) per velocity intervals

Figure 6: Ping RTT characteristics

but also to gain further understanding of the factors affecting observed behaviors.

Here, we in particular consider changepoint detection, and propose an approach which is generic but also well suited to the characteristics of this particular data set. Changepoint detection allows the automated detection of changepoints, and by examining for which metrics the changepoints happen at the same points in time, hints on the underlying relationships between metrics as well as an increased understanding of the reasons for changes in metric values can be obtained.

Changepoint detection has applicability in a wide range of domains, and has been studied extensively. A changepoint for some metric can occur with respect to a change of the location of the mean, variance, or time series autocorrelation of the metric data series. Changepoint detection approaches can be on-line or off-line, where on-line approaches aim to detect a change in a continuously arriving stream of data, and off-line approaches process a complete data set. Different algorithms are available for only detecting the presence or absence of a changepoint, or also estimating the time placement of single or multiple changepoints. A detailed description of methods and applications of changepoint detection is provided in [3]. Here the main interest is the detection, and placement in time, of one changepoint for the location of the mean using an off-line approach.

A. *The Sigmoid-fitting changepoint detection approach*

Based on the wide variety of the distributional properties observed in the metrics present in the data set, and weaknesses of the tested implementations of changepoint detection algorithms such as CUMSUM, a novel approach for changepoint detection based on least-squares non-linear regression of a restricted sigmoid function is proposed. The sigmoid function can be seen as a generalized logistic function, and thus represents two stationary means with a changeover period. The function is in this context expressed as:

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

where Y_0 and Y_1 corresponds to the mean values before and after the changeover, which occurs at time T_c. The S parameter models the shape of the changeover transition and is dependent on the considered time period. Here it is fixed at 1 which, for the considered data, will create a steep slope at the changeover time. By restricting the sigmoid shape parameter S a fairly steep changeover is guaranteed, and fitting is simplified when only three parameters (Y_0, Y_1, T_C) needs to be located. Fitting of the sigmoid parameters is done using non-linear Levenberg-Marquardt (LM) least-squares regression. While regular linear least-squares regression is guaranteed to find the global minimum, non-linear least-squares regression may become stuck at local minima. To avoid the risk of non-global minima being selected, the fitting is performed with 20 runs having uniformly distributed T_c seed values. This multiple

fitting with different starting points has the advantage that the gradient-descent based optimization function employed for the non-linear regression will help to reflect the relative strength of the detected changepoint. A very clean and abrupt change will thus lead to all the 20 regression runs finding the same global changepoint. Weak changes will only result in a single or a few regressions located at that changepoint and the remaining regressions will be located at other minima. This allows for a natural way to perform a post-computation trade off between the true-positive and false-positive rates of the changepoint detection. This is an advantage given the large data amounts, and an aspect which can be problematic for other changepoint detection methods.

Based on the parameters obtained from the LM regression fitting, a changepoint coefficient is calculated for each of the 20 regression runs. The changepoint coefficient captures the relative amount of change that is present in the considered changepoint and is computed as follows:

$$C(T_c) = \frac{|Y_1 - Y_0|}{\min(Y_0, Y_1)} \tag{2}$$

The regression with the maximum C is used as changepoint candidate for the data, and the number of regression runs located around a similar T_c is counted.

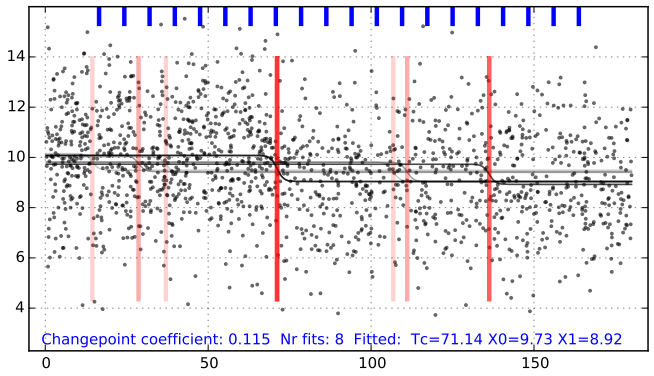


Figure 7: Changepoint detection on synthetic example

A synthetic example is shown in Figure 7. The data in the figure is generated by two normal distributions with parameters $Y_0 = 10, Y_1 = 9, \sigma_{1,2} = 3, T_c = 70$. There were 800 data points generated as shown in Figure 7. Also shown in the figure is the detected changepoint as red bars, the fitted sigmoid functions, the maximum changepoint coefficient and the number of regressions that located the changepoint to within $T_c \pm 0.1$. As can be seen, out of the 20 LM regression runs each with different T_C initial values (blue sticks on top), 8 runs regressed to the same changepoint of 71.14.

As part of ongoing work we are examining approaches to extend the changeover coefficient calculation by including outlier rejection, Box-Cox transformation and data scaling to further improve the detection characteristics over diverse data sets.

B. Cellid changepoint examination

We employ our changepoint detection approach on a subset of the data described earlier in order to highlight its usefulness. For the Stockholm-Karlstad route, 200 frequently occurring cellids out of the 2765 observed on the route were selected and used for further processing. This data set comprised of around 8 million data points, which were grouped into bins according to cellid and individual train passings. The metrics of interest were then averaged over each such cellid - trainpassing event. This resulted in 282166 cellid - trainpassing aggregated data points. These data points were then used for the changepoint detection. In summary, the steps taken are thus: (1) for all metrics, compute aggregated averages per cellid - trainpassing event. For each metric and cellid then: (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 located within 0.1 days of the highest changepoint coefficient. Here, the threshold on the number of fits for a changepoint candidate to be considered valid was set to 5 out of 20.

The changepoint detection was performed for ten of the metrics available in the data set. The selected metrics can be grouped into three distinct categories. The first category consists of metrics that are related to external characteristics that may affect other observed metrics, and include train velocity and the number of active devices. The second category is user-level performance metrics and these include per-link uplink and downlink throughput, and ping RTT. The third category contains radio-related metrics as reported by the modem and includes dbm, RS-SINR, RSRP, RSRQ and Transmit power.

The distribution of the aggregated data over time, and the detected changepoints are shown in Figure 8. The figure shows two example cellids out of the 200 processed cellids, showing one cellid per column. It can be observed that for both cellids, changes were detected for the two topmost external metrics. Although we cannot conclusively determine the cause of the detected changes, we note that the detected change occurs at a time consistent with the start of the Swedish vacation period. As these external metrics are mainly journey-related and have little relation to individual cellid conditions, it could be expected that changepoints are present for many of the cellids. It was found that a considerable majority of all cellids indeed had changepoints detected at similar locations, although more active device changepoints than velocity changepoints were noted.

Looking at the user-level metrics, it can be seen that only one changepoint was detected, for the ping RTT metric of the second cellid. The placement of the changepoint is very close to the changepoint in number of active devices. Consistent with the buffer-bloat discussion in Section IV-B, it can be reasonably assumed that the increase in number of active devices is causing the detected increase in ping RTT. It can be noted that the changepoint was rather weak with 6 close fits. A similar visual pattern for the ping, with fewer low-value pings after the potential changepoint, is also hinted in the data

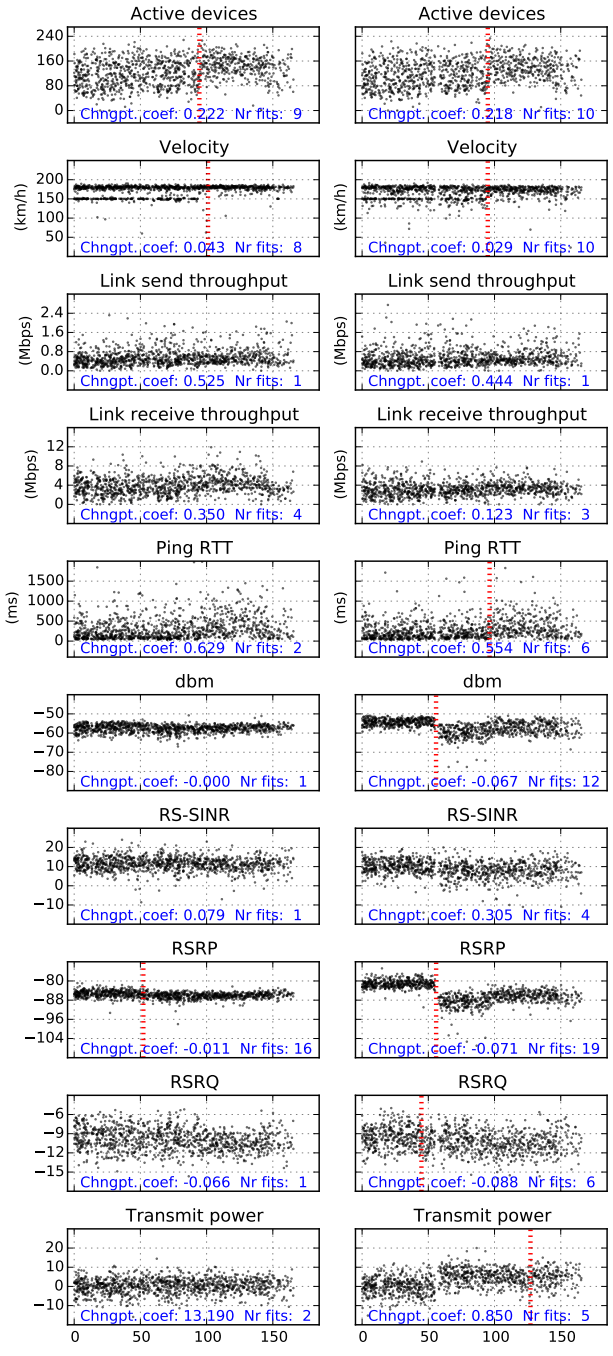


Figure 8: Time evolution of metrics for two example cellids, day since measurement start on x-axis

for the first cellid.

For the final group of radio related metrics, the second cellid shows a distinct pattern. Several of the radio related metrics show a changepoint at the same point in time. It can also be noted that there appears to be a gap in the collected data. This could suggest that this cell was inoperable for a short period of time. When back online again, some aspect of cell operation had changed for the worse, which is reflected in the worse numbers observable for several of the radio metrics. As can be seen, dbm and RSRP changes were easily detected, with high number of fits indicated. A change in RSRQ is detected, but with a slight offset. The somewhat anomalous changepoint detected for transmit power is due to a lack of numerical stability for the changepoint coefficient calculation when the parameters Y_o and Y_1 involve a zero-crossing, as is the case here for the dominant changepoint around day 54. Addressing this issue is part of ongoing work.

VI. Conclusions

In this paper we provide a large-scale analysis of the characteristics of train-based communication systems, where external communication is bundled over multiple cellular links. It is based on an extensive data set gathered over six months from over 7000 journeys within an operational Swedish railway system. The results confirm the strong correlation between the number of users in the system and the observed aggregate throughput. The ping round trip times are also highly dependent on the number of users and increase significantly in the presence of higher load, indicating serious buffer bloating issues in this context. Based on a changepoint method, we examine changes in the collected metrics over time. This to identify possible anomalies in the hardware or environment and to further our understanding of the factors affecting the observed behaviors.

For future work we plan to refine our changepoint detection approach and also hope to complement the current passive measurements with additional active measurements which could provide further insights into the observed behaviors.

Acknowledgments

The authors wish to thank Mats Karlsson and Constanze Deiters for assisting with data collection and processing.

References

[1] M. Alasali and C. Beckman. LTE MIMO performance measurements on trains. In *European conference on Antennas and Propagation (EuCAP)*. IEEE, 2013.

[2] S. Alfredsson et al. Impact of TCP Congestion Control on Bufferbloat in Cellular Networks. In *Proc. IEEE WoWMoM*, 2013.

[3] M. Basseville et al. *Detection of abrupt changes: theory and application*, volume 104. Prentice Hall, Englewood Cliffs, 1993.

[4] J. Calle-Sánchez et al. Long term evolution in high speed railway environments: feasibility and challenges. *Bell Labs Technical Journal*, 18(2):237–253, 2013.

[5] H. Jiang et al. Tackling bufferbloat in 3G/4G networks. In *Proc. ACM Internet Measurement Conference (IMC)*, 2012.

[6] L. Li et al. A measurement study on TCP behaviors in HSPA+ networks on high-speed rails. In *IEEE INFOCOM*, 2015.

[7] Y.-B. Lin, S.-N. Yang, and C.-T. Wu. Improving handover and drop-off performance on high-speed trains with multi-RAT. *Trans. Intelligent Transportation Systems*, 15(6), 2014.

[8] A. Lutu et al. The Good, the Bad and the Implications of Profiling Mobile Broadband Coverage. *Computer Networks*, 2016.

[9] É. Masson, M. Berbineau, and S. Lefebvre. Broadband Internet access on board high speed trains, A technological survey. In *Communication Technologies for Vehicles*. Springer, 2015.

[10] R. Merz et al. Performance of LTE in a high-velocity environment: A measurement study. In *Proc. All Things Cellular (ATC)*. ACM, 2014.

[11] M. K. Müller, M. Taranetz, and M. Rupp. Providing current and future cellular services to high speed trains. *Communications Magazine, IEEE*, 53(10):96–101, 2015.

[12] J. Rodriguez-Pineiro et al. LTE downlink performance in high speed trains. In *IEEE Vehicular Technology Conference (VTC Spring)*, 2015.