Run-time Delays in Indian Railways: is Traffic the Cause?

Saptarshi Ghosh, Sanket Agarwal, Himanshu Srivastava, and Animesh Mukherjee
Department of CSE, Indian Institute of Technology Kharagpur, Kharagpur 721302, India

ABSTRACT

The Indian Railways (IR) has recently come under intense criticism for not being able to efficiently handle the rapidly increasing amounts of traffic, which is manifested in large run-time delays of trains. Motivated by this situation, we perform a statistical analysis of the run-time delays of express trains in IR. We find that trains in some specific geographical regions, such as the Indo-Gangetic plain, regularly experience high delays, indicating high congestion in these regions. However, rather surprisingly, the amount of traffic alone cannot explain the geographical variation of delays in IR. We also identify that the IR time-table contains significant slack, whereby trains running on certain segments frequently traverse the segments in lesser time than is scheduled. To our knowledge, this is the first study on the run-time delays of trains in IR. We also develop a Web-based application for visualizing traffic-flow in IR.

Categories and Subject Descriptors: K.4.1 [Public Policy Issues]: Regulation, Human safety.

Keywords: Indian Railway, traffic, run-time delay.

1. INTRODUCTION

The Indian Railways (IR) is the largest rail-passenger carrier in the world, and has long served as the backbone of the country's transport infrastructure. The increasing demand for transportation in the growing economy of India has resulted in an enormous rise in the volume of IR traffic in recent years. However, it is a commonly voiced opinion that the current IR infrastructure is not capable of efficiently handling this increased traffic [2], which is manifested as regular run-time delays of trains and even frequent accidents [1]. Under these circumstances, a detailed understanding of the problems in IR is essential for adopting effective extension policies in future. and a better planning of the railway budget.

In a recent study [1], we had analysed the traffic flow on the IR network, and found that large amounts of IR traffic are concentrated in some specific geographical regions, such as the Indo-Gangetic plain and central India, which is likely to lead to congestion in these regions. This study [1] considered the traffic according to the official IR time-table (i.e., the ideal or scheduled traffic flow). However, in reality, trains in IR are frequently delayed due to congestion [2]; thus, the *actual* flow of traffic may be significantly different from the scheduled flow. Hence, in the present work, we collect extensive data on the *actual* flow of traffic (trains) in IR, and analyze the statistics of the run-time delays of trains. To the best of our knowledge, this is the first systematic study of the run-time delays of trains in IR.

DEV'13, January 11–12, 2013, Bangalore, India.

2. DATASET

We follow the methods used in our earlier study [1] to construct the dataset. We consider only the 'express' passenger trains in IR, and only the important 'trunk-routes' which connect major stations and are mostly used by express trains. We partition the edges in the IR trunk-route network (see *http://tinyurl.com/trunk-map-ir*) into a set of 54 disjoint 'trunk-segments', where each trunk-segment is the portion of a trunk-route between two junction stations (i.e., stations where multiple trunk-routes converge). To study the traffic and delay patterns in different regions of the country, each trunk-segment is assigned to one of six geographical zones – north, east, west, south, central and the Indo-Gangetic Plain (IGP) – according to its location. Note that we consider IGP (which comprises parts of the north and east zones) as an individual zone in view of the large amounts of traffic (and hence likely congestion) in this region [1].

We collected the delay-status of each express train that uses at least one of the 54 trunk-segments, at each station in its route, from the official website *www.trainenquiry.com*. This site provides daily updates of the *actual* arrival and departure times (which may be different from the scheduled times) of express trains. Data from this website needs to be collected on a daily basis, since on a particular day, only the delays of trains which travelled on the previous and the current days are available. We attempted to collect the delay-data on each day during April–June, 2012, and could obtain the data on 40 distinct days (data on the other days could not be collected due to network issues).

Note that the most challenging aspect of conducting a study on the IR is the unavailability of organized data. For instance, no data on the freight traffic or the traffic-handling capacity of the various trunk-routes are publicly available. We have requested the IR authorities for these data, and are yet to receive a response from them. Hence the dataset gathered by us is possibly the best that can be collected from public sources.

3. ANALYSIS OF RUN-TIME DELAYS

We measure the number of trains using a trunk-segment on each day, from the data collected over 40 days, and then compute the average number of trains using a trunk-segment per day. The run-time delay of a train T on a given trunk-segment is measured as follows. Assume that the end-stations of the segment are s and t, and that the train is travelling from s towards t. The run-time delay of T on this segment is $(d_t - d_s)$ where d_s and d_t are the delays of T at s and t respectively (as obtained from *www.trainenquiry.com*). The delay can be negative as well, if the train is delayed less at t than at s (i.e., if the train takes lesser time to traverse the segment

Segment (highest traffic)	Average traffic	Average delay	Zone	Segment (highest delay)	Average delay	Average traffic	Zone
Delhi-Tundla-Kanpur	84.54	30.75	IGP	Patna-Asansol	42.94	20.74	IGP
Ahmedabad-Surat	80.80	1.92	west	Delhi-Tundla-Kanpur	30.75	84.54	IGP
Delhi-Mathura-Agra	77.40	8.84	IGP	Mughalsarai-Patna	27.69	50.00	IGP
Jhansi-Bina-Bhopal	74.49	6.36	central	Mumbai-Madgaon	22.75	16.51	west
Bhusaval-Manmad-Kalyan	73.11	-2.45	west	Coimbatore-Chennai	21.07	42.11	south

Table 1: The top 5 trunk-segments ranked according to (i) average per-day traffic (on the left), and (ii) average run-time delay of trains on the segment (on the right), along with the geographical zones of the segments. Average delay values are in minutes.

than according to the schedule). The average run-time delay of trains on a segment is computed as the average of delays of all trains using the segment over 40 days. Table 1 shows a ranked list of the top 5 trunk-segments according to (i) the average number of trains using a trunk-segment per day, and (ii) the average run-time delay of trains on a segment. The analysis of the run-time delays led to several interesting observations, which are described below.

High delay in IGP zone: Table 1 shows that 3 out of the top 5 segments having highest run-time delays are from the IGP zone. In fact, 13 out of the 15 segments showing the highest average delays are from the IGP zone (full list not shown for lack of space), which indicates exceptionally high congestion in this zone.

Low correlation between traffic and delay: Table 1 also shows that the segments having the highest per-day traffic and the ones having highest delay show *low* overlap. In fact, there are examples where segments having very high daily traffic show low delays and vice-versa. For instance, the Bhusaval-Kalyan segment has an average daily traffic as high as 73.11, but shows a negative delay, whereas the Mumbai-Madgaon segment has a low average daily traffic of 16.51 but a relatively high delay of 22.75 minutes (see Table 1). Overall, considering all 54 segments, the average per-day traffic and the average delay for a segment has a low positive Pearson correlation coefficient of 0.13.

Delay at different times of day: We also observed the temporal variation of run-time delays over a day, for each segment (results omitted for lack of space). We divided a day into 12 slots, each of duration 2 hours (00:00 – 02:00, 02:00 – 04:00, ...) and measured the average run-time delay of trains which use a given segment during each slot. We found that several segments show high delays during the late hours at night, which signifies high congestion during these hours in several regions of the country.

Segments with negative delay: Out of the 54 segments considered, 13 exhibited negative delays on average, i.e., most trains traverse these segments in lesser time than is scheduled in the time-table. This indicates that the IR schedule has unnecessary *slack* – possibly the time required by trains to traverse these segments has reduced over the years due to improved technology or increase in resources such as parallel tracks, but the IR time-table has not been updated to reflect this fact.

4. VISUALISATION OF IR TRAFFIC

Our studies indicate that the amount of IR traffic varies widely from segment to segment; further, traffic on a given segment also shows large variation with the time of day. We have developed a Web-based application system for visualizing the amount of traffic on different trunk-segments at specific hours of the day. The application shows the trunk-segments on a map of the country (Fig. 1). The amount of traffic on a certain segment at a particular hour of day is graphically shown as a line on the segment, where the thickness of the line is proportional to the number of trains using the segment at the given hour. The application can be used to visually monitor flow of traffic through the segments throughout a day – for this, the flow of traffic is *simulated* (using computer simulations) according to the real data of the traffic flow on the previous day

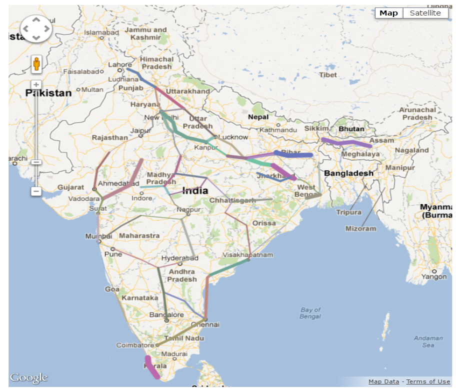

Figure 1: Screen-shot of the application showing IR traffic in different trunk-segments at 23:00 hours. The thicker lines indicate segments having higher traffic during this hour.

(as crawled from the IR website). A suitable scaling of time can also be applied, e.g., the flow of traffic over an hour can be simulated and visualized over 1 minute. Such an application can be used by the IR authorities to systematically study the traffic scenario in different IR segments over a sizeable number of days, in order to identify trends in the flow of traffic.

5. CONCLUDING DISCUSSION

To our knowledge, this is the first systematic study on the run-time delays of trains in IR. The study identified several problems in IR, such as high run-time delays (indicating high congestion) in the IGP zone, high delays in many segments during the late hours of the night, presence of significant slack in the time-table for some of the segments, and so on. Most importantly, we, rather surprisingly, found that the amount of traffic does not seem to be the only factor contributing towards this delay. In some cases, high delays are correlated with high traffic, e.g., in case of segments in the IGP, or in several segments during the late hours at night. However, delay is less correlated (or, at times, negatively correlated) with traffic for some other geographical regions.

This indicates that there are likely to be other factors which contribute towards the run-time delay and congestion in IR. Some likely factors are varying traffic-handling capacities of the segments (e.g., number of parallel tracks), difference in signalling systems, the effects of freight traffic (which may be skewed towards some segments which show high delays in spite of having low passenger traffic), and human factors such as poor administration. These issues need to be investigated further by the IR authorities. We also hope that the IR authorities will make data on track capacities and freight traffic public in the near future, to enable more detailed investigation of the traffic and delay scenario.

6. REFERENCES

[1] S. Ghosh, A. Banerjee, and N. Ganguly. Some Insights on the Recent Spate of Accidents in Indian Railways. *Physica A*, 391(9), 2012.

[2] Have India's clogged rail lines stalled progress? `http://tinyurl.com/clogged-ir`, June 2010.