Modeling Delay Propagation in Subway Systems with Bayesian Hierarchical Models

Shayan Nazemi[1], Aurélie Labbe[1], Stefan Steiner[2], Martin Trépanier[3], and Léo Belzile[1]

[1]*Department of Decision Sciences, HEC Montréal, Montréal, Québec, Canada*
[2]*Department of Mathematical and Industrial Engineering, Polytechnique Montréal, Montréal, Québec, Canada*
[3]*Department of Statistics and Actuarial Sciences, University of Waterloo, Waterloo, Canada*

Abstract

Disruptions are an unavoidable aspect of transportation systems, often occurring unpredictably and varying in duration. These events impact network dynamics by affecting the punctuality of both the disrupted train and preceding trains. Accurate information on waiting and journey times is essential for passengers during the post-disruption period specially in systems without a predefined timetable. This study investigates delay propagation in a subway system immediately after a disruption is resolved by modeling dwell and running times through a combination of mixture analysis and Bayesian hierarchical models. Our proposed model uses the track occupation data from Montreal's metro system, and leverages incident logs in the network to separate disruptions from normal operations, enabling analysis of process times and delay propagation.

Motivation

Operational procedures in metro systems are determined by the duration trains remain at stations, known as dwell times, and the travel time between stations, referred to as running times, which we collectively refer to them as process times of trains within the network. Key factors in timetable construction include dwell times at stations, running times across different sections, and the headway between consecutive trains. Passenger journey times can be broken down into the sum of all the dwell and running times between an origin and a destination. Modeling the joint distribution of dwell and running times are essential in accurate estimation of journey times, however, process time modeling is challenging due to the complex spatio-temporal correlations between dwell and running times. Delays in transportation networks lead to longer process times, and their impacts are not necessarily contained to a single train, station or line. During disruptions, trains often remain at stations until the issue is resolved, causing passenger buildup at stations further along the route. As a result, the primary delay of a train can be extended to the rest of its journey (self-propagation) causing longer dwell processes after the resolution of incident. Headway constraints force following trains to stop at preceding stations, affecting the punctuality of subsequent trains (backward propagation). In cases of major disruptions, the closed-loop nature of metro lines can also impact train operations in the opposite direction. Depending on multiple factors such as location, duration and type of the delay, process times are longer during the incident recovery period, and they tend to go back to their steady-state as time progresses. Delays can have different cascading implications on a network level (Li, Guo, et al. 2021), and disruption management plans apply measures to resolve traffic and prevent delay propagation to the rest of the network (Zilko et al. 2016). Running times exhibit low variability and remain largely unaffected by external factors. The primary source of uncertainty in process time modeling thus arises from the high variability of dwell times, which comes from the time-varying nature of passenger flow. Our research seeks to model the distribution of dwell and running times during post-disruption periods using a hierarchical model that accounts for the network's structural dependencies.

Case Study

This study is motivated by the sensory track occupation data provided by the Société de transport de Montréal (STM) for all four metro lines in Montréal, Québec, Canada. These lines collectively include 73 stations and cover a total distance of 69 kilometers. Train operations do not adhere to a fixed timetable, with service headways ranging from 2 to 10 minutes, varying between peak and off-peak hours as well as across weekdays and weekends. Service hours extend from 5:30 A.M. to 1:00 A.M., with slight variations on weekends. The dataset spans the entire year of 2018, containing over 126 million track occupation records, with an average of approximately 240 train passes per direction on weekdays and 150 on weekends. Additionally, disruption logs from the same period, including timestamps, durations, locations, and causes of incidents, were provided and used to filter disruption records and analyze delay propagation in track occupation dataset. There are 1,055 disruption records for the entire year of 2018 with an average duration of 10.4 minutes. We synchronize the incident reports with the track occupation records of trains in the network to examine the impact of factors such as location, distance to the incident, rush hour conditions, and more.

Literature Review

Minimum dwell time depends on boarding, alighting, and onboard passengers, while the random component's conditional distribution varies with passenger flow levels (Cornet et al. 2019). Several studies have explored the statistical properties of dwell times indicating that they follow a heavy-tailed distribution (Li, Goverde, et al. 2014; Pang et al. 2023). Obtaining and utilizing passenger flow data for real-time predictions is challenging, as this information is often unavailable in real-time (Li, Daamen, et al. 2016). Several data-driven approaches have been developed in recent years to study delay propagation in railway network systems, focusing on the modeling of process times (Kecman et al. 2015; Li, Daamen, et al. 2016; Cornet et al. 2019; Li, Guo, et al. 2021). Kecman et al. (2015) estimate railway process times based on different statistical learning techniques such as robust linear regression and random forests, and showed the effectiveness of covariates such as headway, peak hours, and arrival delay in process time estimation. Li, Daamen, et al. (2016) propose a regression model to estimate dwell times at short stops during peak hours and a non-parametric regression model for off-peak hours. Cornet et al. (2019) model dwell times using passenger flow data, dividing it into a deterministic minimum dwell time and a random component.

Although these methods improve the empirical estimation of delays and process times, they do not take into account the complex network structure of railway systems into their modeling, which results in complicated dependence between operation events such as delays, control actions, and process times. Previous studies in delay propagation analysis has explored probabilistic networks to model these dynamics (Bearfield et al. n.d.; Sun et al. 2015; Ulak et al. 2020), with the majority of solutions relying on the Markov property in the networks (Li, Guo, et al. 2021; Corman et al. 2018). The advantage of probabilistic graphical models lies in their ability to update their belief about future events as information is received. For example, Li, Guo, et al. (2021) propose a conditional Bayesian model that could address different delay propagation scenarios (self, backward, and cross-line) based on the incremental running and dwell times. Corman et al. (2018) use delay events as nodes in a Bayesian network to analyze delay dynamics and assess its uncertainty with incoming traffic information. Ge et al. (2024) model delays through operational control actions while relaxing the Markov assumption. They analyze how station and section control actions influence downstream delays, considering the states of the two preceding trains. Their studies showed that section control actions have higher intensity influence on train delays with respect to station control actions.

Methodology

We model the transitional dynamics following an incident through a Bayesian hierarchical model that captures the autocorrelation of dwell times across consecutive trains (temporal dependence) and stations (spatial dependence). We focus on the dwell (running) times starting at a station (section) level, specifying the dependence within a single metro line in one direction using an autoregressive process with time-varying parameters that smoothly vary across stations. Given the heavy-tailed nature of dwell time distributions and the varying operational adjustments after incident resolution, we employ log-normal mixture models to cluster dwell processes. These clusters are informed by temporal and spatial factors, as well as external influences such as rush hour conditions, time elapsed since incident resolution, distance from the disruption, and whether the incident was classified as prolonged. Our analysis is a two-step procedure where, in step 1, a spatio-temporal mixture model is fitted to classify observations into different dwell operation procedures and, in step 2, a log-normal hierarchical Bayesian spatio-temporal model is used for each category to predict the process time.

Step 1: Let $Y_{i,s}$ denote the dwell time of the ith train at station s. Given a vector of external factors $\mathbf{x}_{i,s} \in \mathbb{R}^p$, the distribution of the dwell process follows a log-normal mixture model, where the mean specification is a linear combination of dwell time of the previous train at the same station (temporal auto-regression), $Y_{i-1,s}$, and dwell time of the same train at the previous station $s-1$ (spatial auto-regression), $Y_{i,s-1}$, as well the factors related to the disruption ($\mathbf{x}_{i,s}$). If we consider the random variable $Z_{i,s} = \log Y_{i,s}$, then:

$$
\begin{aligned}
f(Z_{i,s}) &= \sum_{k=1}^{K} \pi_k^* \Phi(\mu_{i,s,k}^*, \sigma_{s,k}^{2*}) \\
\mu_{i,s,k}^* &= \beta_0^* + \phi_s^* Z_{i,s-1} + \phi_t^* Z_{i-1,s} + \boldsymbol{\beta}_k^{*\top} \mathbf{x}_{i,s}
\end{aligned} \tag{1}
$$

where K is the number of mixtures, π_k^* is the mixing probability of mixture k, and $\Phi(\cdot)$ is the normal distribution function. The asterisk ($*$) superscript in equation (1) is used to distinguish between the estimated parameters in the mixture analysis (step 1) from those in the Bayesian hierarchical model (step 2). We use four covariates in our analysis ($p = 4$): $\mathbf{x}_{i,s} = [x_{1,i,s}, x_{2,i,s}, x_{3,i,s}, x_{4,i,s}]^\top$, where $x_{1,i,s}$ is a rush hour indicator, $x_{2,i,s}$ is the number of minutes since the resolution of incident, $x_{3,i,s}$ indicates whether the incident was considered as a major disruption, and $x_{4,i,s}$ is the distance to the incident location in terms of number of stations for the i-th sample.

The mixture assignments are determined using the Expectation-Maximization (EM) algorithm, which iteratively updates the probability that each observation belongs to a specific mixture component. While the mixture variances are assumed to be independent of the covariates, preliminary results indicate that the variability of the dwell process increases during rush hours and major disruptions.

Step 2: After clustering, each component is modeled separately using a Bayesian hierarchical model. The mean specification accounts for spatio-temporal dwell processes and external factors ($\mathbf{x}_{i,s}$), but unlike the initial step, the variance specification is also a function of covariates, namely the rush hour indicator ($x_{1,i,s}$) and the classification of the incident as a major disruption ($x_{3,i,s}$). A log-normal distribution is assumed for the dwell processes, as they better account for extreme values than a normal distribution.

For mixture k, the data generating process is:

$$
\begin{aligned}
Y_{i,s} &\sim \text{Log-Normal}(\mu_{i,s,k}, \sigma_{i,s,k}^2) \\
\mu_{i,s,k} &= \beta_{0,k} + \phi_{s,k} Z_{i,s-1} + \phi_{t,k} Z_{i-1,s} + \boldsymbol{\beta}_k^\top \mathbf{x}_{i,s} \\
\sigma_{i,s,k} &= \exp(\theta_{0,k} + \theta_{1,k} x_{1,i,s} + \theta_{2,k} x_{3,i,s})
\end{aligned} \tag{2}
$$

where we assign normal priors to $\boldsymbol{\beta}, \boldsymbol{\theta}$, and $\boldsymbol{\phi}$.

Preliminary Results

To showcase the model, we focus on post-disruption period consisting of up to twice the time disruption length from incident resolution leading to a sample of 1,088 dwell times at Peel station. We use information criteria to determine the optimal number of mixture components in step 1, which is $K = 3$. The results show a mean absolute error (MAE) for predicted vs. realized dwell times of 6.65 seconds and a mean absolute percentage error (MAPE) of 8.65% across all three mixtures based on five-fold cross validation. The three mixtures shown in Figure 1, correspond to different dwell processes, reflecting distinct train operation procedures (short, normal, and long dwell processes). The parameter estimates indicate positive effects of dwell processes over time and space in both short and normal modes of operation, with $\hat{\phi}_{s,1} = 0.110, \hat{\phi}_{t,1} = 0.161$, and $\hat{\phi}_{s,2} = 0.086, \hat{\phi}_{t,1} = 0.032$; however, there seems to be no significant effect for the long dwell times in the third mixture. The mixture associated with long dwell processes shows the greatest impact from rush hours, with a mean dwell time 3.26 times higher than periods outside of rush hour. The model demonstrates higher accuracy for short and normal dwell operations but tends to underestimate very long dwell processes. Further model tuning is needed to better capture long dwell times. The underestimation may be due to the high variability and heavy-tailed nature of extreme dwell times, where the log-normal distribution struggles to fit exceptionally large values.[1]

[1]Rephrasing assistance provided by ChatGPT, a language model developed by OpenAI

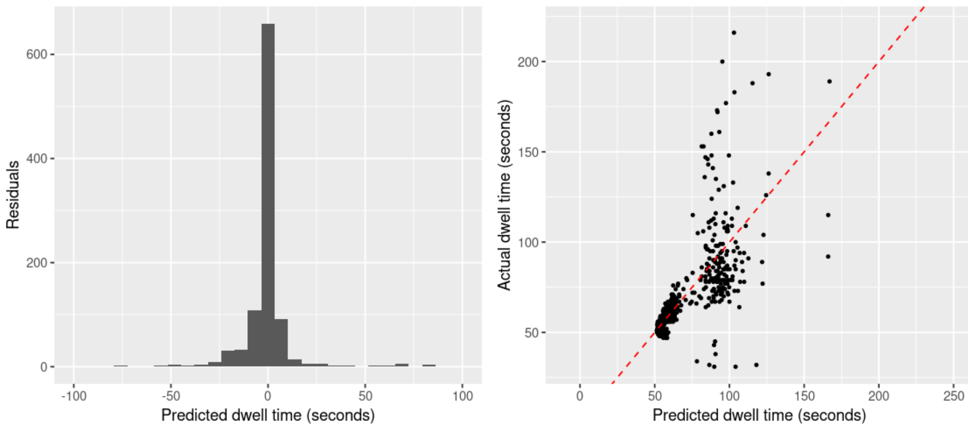

Figure 1: Histogram of model residuals (left) and predicted vs. realized dwell times (right) for the post-incident records at Peel station.

References

Bearfield, George, Anna Holloway, and William Marsh (n.d.). “Change and safety: decision-making from data”. In: *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* (). DOI: 10.1177/0954409713498381.

Corman, Francesco and Pavle Kecman (2018). “Stochastic prediction of train delays in real-time using Bayesian networks”. In: *Transportation Research Part C: Emerging Technologies*. DOI: 10.3929/ethz-b-000281218.

Cornet, Sélim et al. (2019). “Methods for quantitative assessment of passenger flow influence on train dwell time in dense traffic areas”. In: *Transportation Research Part C: Emerging Technologies* 227.6, pp. 704–714. DOI: 10.1016/j.trc.2019.05.008.

Ge, Xuekai et al. (2024). “Modelling the cascading effects of train delay patterns and inter-train control actions with Bayesian networks”. In: *International Journal of Rail Transportation*. DOI: 10.1080/23248378.2023.2194304.

Kecman, Pavle and Rob M.P. Goverde (2015). “Predictive modelling of running and dwell times in railway traffic”. In: *Public Transport*. DOI: 10.1007/s12469-015-0106-7.

Li, Boyu, Ting Guo, et al. (2021). “Delay propagation in large railway networks with data-driven Bayesian modeling”. In: 2675 (11), pp. 472–485. DOI: 10.1177/03611981211018471.

Li, D., R. M.P. Goverde, et al. (2014). “Train dwell time distributions at short stop stations”. In: *2014 17th IEEE International Conference on Intelligent Transportation Systems, ITSC 2014*, pp. 2410–2415. DOI: 10.1109/ITSC.2014.6958076.

Li, Dewei, Winnie Daamen, and Rob M. P. Goverde (2016). “Estimation of train dwell time at short stops based on track occupation event data: A study at a Dutch railway station”. In: *Journal of Advanced Transportation* 50.5, pp. 877–896. DOI: 10.1002/atr.1380.

Pang, Zishuai et al. (2023). “Dynamic train dwell time forecasting: a hybrid approach to address the influence of passenger flow fluctuations”. In: *Railway Engineering Science* 31 (4), pp. 351–369. DOI: 10.1007/s40534-023-00311-7.

Sun, Lijun et al. (2015). “An integrated Bayesian approach for passenger flow assignment in metro networks”. In: *Transportation Research Part C: Emerging Technologies* 52, pp. 116–131. DOI: 10.1016/j.trc.2015.01.001.

Ulak, Mehmet Baran, Anil Yazici, and Yun Zhang (2020). “Analyzing network-wide patterns of rail transit delays using Bayesian network learning”. In: *Transportation Research Part C: Emerging Technologies*. DOI: 10.1016/j.trc.2020.102749.

Zilko, Aurelius A., Dorota Kurowicka, and Rob M.P. Goverde (2016). “Modeling railway disruption lengths with copula Bayesian networks”. In: *Transportation Research Part C: Emerging Technologies*. DOI: 10.1016/j.trc.2016.04.018.