Ping Huang J. Lessan Chao Wen Qiyuan Peng Liping Fu Li Li Xinyue Xu

**A Bayesian network model to predict the effects of interruptions on train operations**

Transportation Research Part C: Emerging Technologies · May 2020 vol. 114 pp. 338-358

DOI: 10.1016/j.trc.2020.02.021

Abstract

Based on the Bayesian network (BN) paradigm, we propose a hybrid model to predict the three main consequences of disruptions and disturbances during train operations, namely, the primary delay (L), the number of affected trains (N), and the total delay times (T). To obtain an effective BN structure, we first analyze the dependencies of the involved factors on each station and among adjacent stations, given domain knowledge and expertise about operational characteristics. We then put forward four candidate BN structures, integrating expert knowledge, the interdependencies learned from real-world data, and real-time prediction and operational requirements. Next, we train the candidate structures based on a 5-fold cross-validation method, using the operational data from Wuhan-Guangzhou (W-G) and Xiamen-Shenzhen (X-S) high-speed railway (HSR) lines in China. The best performing structure is nominated to predict the consequences of disruptions and disturbances in the two HSR lines. Comparisons results show that the proposed model outperforms three other commonly used predictive models, reaching an average prediction accuracy of 96.6%, 74.8%, and 91.0% on the W-G HSR line, and 94.8%, 91.1%, and 87.9% on the X-S HSR line for variables L, N, and T, respectively.

Основываясь на парадигме байесовской сети (БС), мы предлагаем гибридную модель для прогнозирования трех основных последствий сбоев и возмущений во время движения поездов, а именно: первичной задержки (L), числа затронутых поездов (N) и общего время задержки (T). Чтобы получить эффективную структуру BN, мы сначала анализируем зависимости вовлеченных факторов на каждой станции и между соседними станциями, учитывая предметные знания и опыт в отношении рабочих характеристик. Затем мы выдвигаем четыре возможных структуры BN, объединив экспертные знания, взаимозависимости, извлеченные из реальных данных, прогноз реального времени и эксплуатационные требования. Затем мы обучаем структуры-кандидаты на основе метода 5-кратной перекрестной проверки, используя рабочие данные линий высокоскоростной железной дороги (HSR) Ухань-Гуанчжоу (WG) и Сямынь-Шэньчжэнь (X-S) в Китае. Наилучшая структура назначается для прогнозирования последствий сбоев и нарушений на двух линиях ВСМ. Результаты сравнения показывают, что предложенная модель превосходит три другие широко используемые прогностические модели, достигая средней точности прогнозирования 96,6%, 74,8% и 91,0% на линии WG HSR и 94,8%, 91,1% и 87,9% на линии X-S HSR для переменных L, N и T соответственно.