

Mining and Learning from Railway Safety Data with Graphs and Tensors

Deliverables and Reporting Requirements for UTC Grants Awarded in 2023 (June 2023)

Exhibit D

Research Project Requirement Template

Recipient/Grant (Contract) Number: University of Texas Rio Grande Valley (UTRGV), University of California Riverside (UCR)/Grant No. 69A3552348340

Center Name: University Transportation Center for Railway Safety (UTCRS)

Research Priority: Promoting Safety

Principal Investigator(s): Dr. Jia Chen (PI, UCR), Dr. Evangelos Papalexakis (Co-PI, UCR)

Project Partners: University of Nebraska Lincoln (UNL), University of South Carolina (UofSC), University of Texas Rio Grande Valley (UTRGV).

Research Project Funding: \$100,000 (Federal), \$50,000 (Non-Federal Cost Share)

Project Start and End Date: 06/01/2023 to 08/31/2024

Project Description: Railway systems are very complex pieces of cyberinfrastructure, interfacing with a number of transportation agents and other pieces of cyberinfrastructure. For instance, a railway crossing includes interactions between the railway system and a traffic intersection. Such a rich ecosystem of interactions among heterogeneous agents poses fascinating research challenges in modeling railway systems with data and conducting data-driven railway crossing safety assessment. In this project, we propose to leverage and extend powerful tensor and graph mining methods which can extract “needles in the haystack” within the abundance of collected data and produce actionable insights to stakeholders in order to better understand emerging accident patterns from historical data, identify underlying similarities in such patterns, towards ultimately reducing the number of accidents.

US DOT Priorities: This project aligns with the following USDOT strategic goals, as established in the USDOT Strategic Plan for FY2022-FY2026: **(a) Safety:** The project directly addresses a major safety concern, railroad grade crossings, which has been identified by the USDOT and the FRA as a major safety concern. **(b) Economic Strength:** Rail accidents cause major economic losses each year, which can eventually be reduced through adoption of the systems developed in this project. **(c) Transformation:** This project addresses the DOT goal of new projects applying novel data approaches based on artificial intelligence to transportation problems. It perfectly dovetails with USDOT’s “transformative vision of a fully connected, integrated, accessible, and interoperative multimodal system of systems” as it seeks to represent data about railroad accidents, an inherently complex interconnected system of systems using graph mining and learning methods. This project aims to make public railroad accident data accessible to human analysts by extracting actionable insights and “needles in the haystack” from large amounts of interconnected multimodal railroad data. The project aligns perfectly with main USDOT objectives for “a fully connected, integrated, accessible, and interoperative multimodal system of systems” which “provides for a critical multimodal transportation knowledge base outside of the USDOT”.

Outputs: We expect to have the following results by the end of the project period:

- a. Research publications, targeted to top-tier data science, machine learning, and artificial intelligence venues.

- b. Publicly available source code for the methods developed. Typically, each publication will be accompanied by a link to publicly available source code on a widely used repository such as GitHub.

Outcomes/Impacts: The methods developed in this project have wide applicability beyond the immediate problem being analyzed and will contribute to the artificial intelligence and machine learning research community as well as the rail industry. It is expected that the results will inform decision making in rail operations and rail policy.

Final Research Report: Upon completion of the project, a URL link to the final report will be provided.