Passenger behavior analysis at Stade de France using GMM model

Minh Duy Nguyen, *ModIA student student, INSA Toulouse & data scientist apprentice at Transilien SNCF-Voyageurs*

Supervisor: **Rémi Coulaud**, *Manager at DataLab' MTA, Transilien SNCF-Voyageurs*

January 2024

SNCF\Transilien\MTA\DataLab'

Transilien is the commuter train network of SNCF Voyageurs, mainly serving stations in the Île-de-France region. Within Transilien, the **Mass Transit Academy (MTA)** promotes train operations in dense areas. I work for the **DataLab' Mass Transit**, which is the innovation laboratory responsible for enhancing passenger and train flow data in the Paris suburbs.

The ASE Project

Context

Study of passenger behavior on the days of the latest music events and international sports competitions at the Stade de France (SDF) prior to the 2024 Olympic Games.

Scope of ASE

Site: RER B, RER D stations and area around SDF
Time: July 2023 – Dec 2023

Fig.1. Studied events. Left: Rugby WC schedule. Right: concerts at SDF since July 2023 (Source: https://www.stadefrance.com/fr)

Objectives

- Understand traveler behavior to prepare the transport plan for Olympics Paris 2024.
- Compare it with concert-type events.

Methodology

Data : count data on the number of passengers entering and leaving the station, using specialized counting cameras.

First analysis

- Plot the time series of passenger data

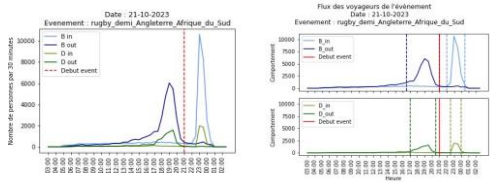


Fig.2. Number of passengers by time on Rugby WC semi-final day at 2 stations in each entry/exit direction. Blue lines represent flows in station RER B and green ones for RER D. Peak-time frames are located between the dotted lines. We observe that on match day, passenger numbers increased significantly.

Calibration of peak-time frames

Use of Gaussian Mixture Model (GMM) → Approximate this time series data → Determine the pe-ak times when fans arrive/ leave the stadium (the period to which 95% of fans turn).

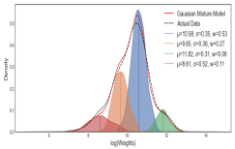


Fig.3. example of Gaussian approximation of a time series (Source: Wikipedia)

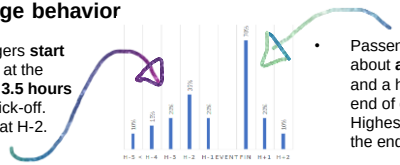
Study of average behavior of two kinds of event

- Find the average performance of time series during peak time frames for each type of event (rugby or concert).
- Perform statistical test for overlap of these two average time series

Results

Average behavior

Passengers **start arriving** at the stadium **3.5 hours** before kick-off. Highest at H-2.



- Passengers **depart** about **an hour** and a half after the end of event. Highest right after the end.

Fig.4. Average passenger behavior for an event at the Stade de France.

Comparing two types of event

- Leaving stream is the same for both types.
- Arriving stream is different: Concert-goers tend to arrive earlier and more slowly

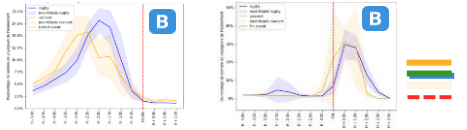


Fig.5. Average behavior of passengers during peak time. Left: Average stream of people arriving at the stadium before kick-off time. Right: Average stream of people leaving the stadium

Conclusion

The flow of people going to the stadium for a sportive event is very large and overwhelming → necessary to plan appropriately for the Paris Olympic Games in 2024.

Promote the upcoming project to model the flow of people moving around the St-ade de France. (Fig.6)

Fig.6. Simulation of passenger flows leaving the Stade de France