Funded by the European Union
NextGenerationEU

SUMMIT
SOFIA UNIVERSITY MARKING MOMENTUM FOR INNOVATION AND TECHNOLOGICAL TRANSFER

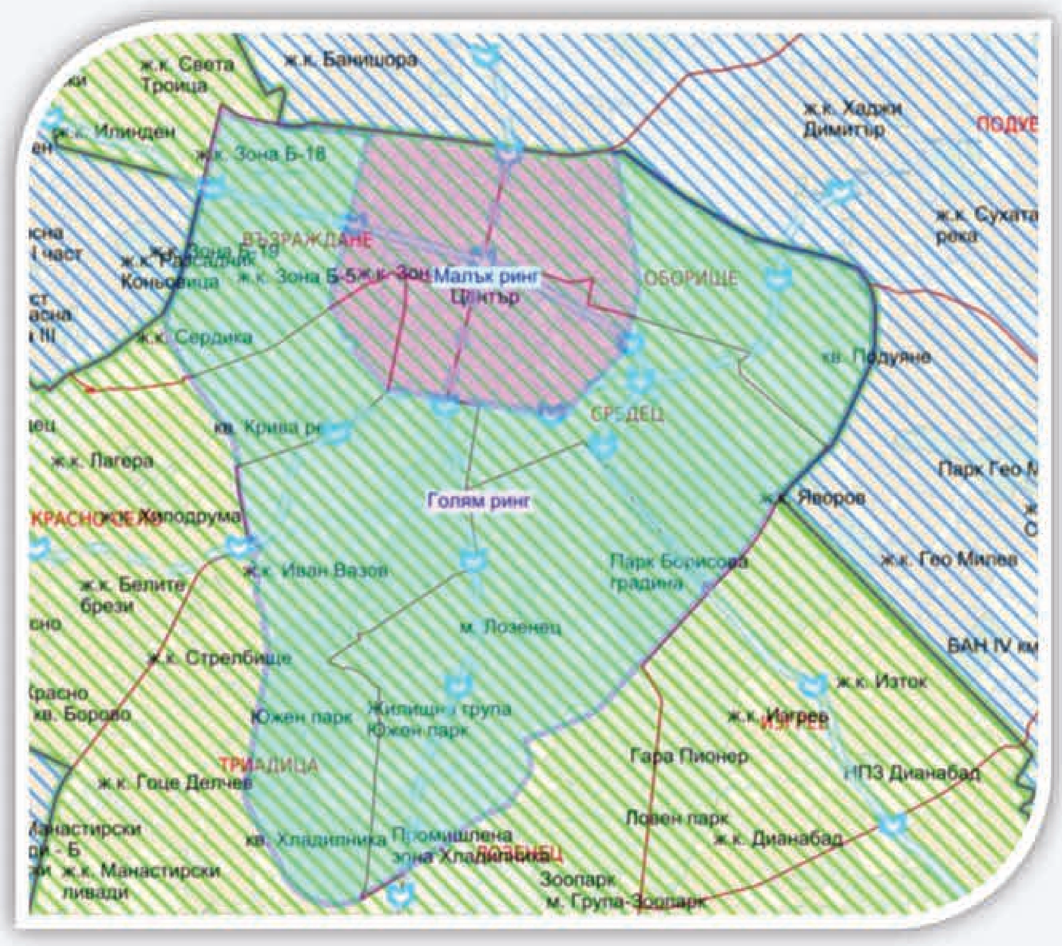
SOFIA UNIVERSITY -
MARKING MOMENTUM
FOR INNOVATION AND TECHNOLOGICAL TRANSFER

National Recovery and Resilience Plan
OF THE REPUBLIC OF BULGARIA

research group
Investigating the effect of urban mobility innovation using mode-switching time series methods
research area: Social Sciences (Political Sciences, law, educational sciences)

Head of the research group:
Assoc. Prof. Dr. Boyan Lomev
Members of the group:
Assoc. Prof. Dr. Marian Milev
Assoc. Prof. Dr. Diana Netova
Assist. Prof. Dr. Teodora Varbanova
Assist. Prof. Dr. Savina Mihaylova
Boyan Davidkov
Konstantin Georgiev
Slavena Dudrenova
Katya Nikolova

Markov Regime-Switching Modeling of Passenger Flow in a Suburban Area Based on Anonymized Telecommunications Data



INTRODUCTION

Urban travel dynamics usually demonstrate a mix of sustainable and changing behavioral patterns that reflect local specifics, such as habits of residents, public holidays, traffic density, seasonality, and weather. This project is focused 0n an original approach for modeling the passenger flow to and from subway stops in suburb an areas of cities based on an anonymized and statistically aggregated tele com operator dataset. The primary data source analyzed in this study was obtained through a collaboration agreement with one of the three main mobile network operators in Bulgaria and contains information about trips from the "Manastirski Livadi" residential complex, located in the suburbs of the city of Sofia, to several different destinations in the capital city. The data are grouped into three indepen dent clusters of travelers:

(1) within the "Manastirski Livadi" residential complex,
(2) to destinations that are accessible by subway or tram, and
(3) to destinations to which there is no convenient "green" public transport.

For each of these three groups, we developed Markov regime-switching models suitable for analyzing passenger flow behavior. We tested the model for different clusters of travelers, specifically in the city of Sofia. The results obtained can be used to create optimal routes and timetables for public transport in suburban areas and to analyze and improve suburban route networks in other regions.

PROJECT GUIDELINES

Over 70% of EU citizens live in urban areas, which generate 23% of all greenhouse gas emissions from transport. The European Green Deal aims to make Europe the first climate-neutral continent by reducing greenhouse gas emissions (at least 55% by 2030 compared to 1990 and by 90% by 2050). The New EU Urban Mobility Framework (Strasburg, 14.12.2021 COM (2021)811final) proposes measures to encourage EU Member States to develop urban transport systems that help address urban mobility challenges, such as air pollution and increase the share of sustainable modes of transport (in particular public transport and active mobility).

Traffic in Sofia has an impact on air pollution and is in third place in terms of pollution with fine dust particles. The number of vehicles registered in Sofia is increasing daily. There are 550-600 cars per 1000 inhabitants. In 2017, a quarter of all vehicles were over 20 years old and a further 60% between 10 and 20 years old. In 2020 Sofia becomes the first Bulgarian city to win funding under the "INNOAIR" project, supported by EC program Urban Innovative Actions that encourages local authorities to search for, test and implement innovative solutions. The Faculty of Economics and Business Administration, Sofia University "St. Kliment Ohridski" participates as a partner and a major part of the team that works according to the present scientific study. Through INNOAIR, "green public transport on demand" was introduced, which is one of a series of measures taken in Sofia to achieve sustainable urban mobility.

Sofia implements various measures to achieve sustainable urban transport, such as low-emission zones, whose effect on air quality in Sofia is precisely the object of research in the scientific study.

METHODOLOGY

The daily and hourly time series used for examining passenger flows of the three zones exhibit two substantially distinct regimes with different intensities, thus characterizing the regular and vacation time mobility of citizens. Consequently, we have two regimes (states) to which we apply Markov Regime-Switching Models. The time series belonging to dissimilar states are modeled by the ARX process with specific structure and variability. We consider univariate ARX (0) and ARX (1) models of daily and hourly data. The exogenous variable x_t for the daily data is the average number of passengers for a day of the week. In the case of hourly data, we have two exogenous variables : the average number of passengers for a specific hour in the day of the week and the "feel like" temperature for the respective hour.

RESULTS

It is important to explore a predictive approach that can be effectively adapted to data with different frequencies and different behavioral patterns. The empirical results and travel demand analysis results are given in two subsections (Subsect. 3.1: Daily Models and Subsect. 3.2: Hourly Models). With the hourly time series models, we are addressing changing behavioral patterns for both high and low traffic regimes in the next day period, and with the daily time series models, we can cover the big picture of changing behavioral patterns governed by local specifics, such as public holidays and seasonality.

Two different regimes of trips can be defined for the period subjec t of the analysis (April 13th to August 31st, 2022): regime 1 (normal activity) and regime 2 (vacation activity). The estimation set contains the April 13th to August 24th period and the test set is the last week of August. Generally, ARX (0) and ARX (1) models are estimated for every time series.
The best estimable models for each zone according to the information criteria are presented below. For Zone 1 daily data, only the ARX (0) model was estimated properly.
Transition matrices for daily data are presented in Table 1.

Zone	State	Pil	Pi2
	state 1, i = 1 Normal activity	0.988	0.012
	state 2, i = 2 Vacation activity	0.063	0.937
2	state 1, i = 1 Normal activity	0.982	0.018
	state 2, i = 2 Vacation activity	0.080	0.920
3	state 1, i = 1 Normal activity	0.991	0.009
	state 2, i = 2 Vacation activity	0.046	0.954

For all zones, we see that the probability for the process to stay **in** the initial state is much higher than the probability of switching. The transition from "vacation activity" to "normal activity" has a slightly higher plausibility. This corresponds to the commonsense intuition that off-duty days are much fewer in comparison with working days.

The only insignificant estimated parameter is the constant term for State 1 in Zone 3, which means that the probability of having zero or negative values is high (its value is 37.5617). The constant in autoregressive models affects a more accurate fit between the model and the data but is not relevant to modeling the correlation structure of the time series. In this sense, taking zero or negative values of the constant for Zone 3 (the constants for Zones 1 and 2 are negative) is not essential for the adequacy of the model. The AR(1) coefficients for all zones and states are positive and correspond to day-to day dependences on passengers' behavior, which is mainly predetermined by seasonal factors.

CONCLUSION

In our research project, we investigated passenger flow dynamics utilizing CDR datasets produced by one of Bulgaria's main mobile network operators. At the core of our research study was the aim of addressing the incapacity of cities to provide optimal public transport routes and schedules that meet citizen demands and needs for public transport services.
This study covers two subperiods with distinct dynamics in their time series representing the transport habits of Sofia residents. Many economic time series exhibit instabilities, often multiple times over their process history . Such is the case with the time series depicting transportation habits among Sofia residents. Under these conditions, we apply a Markov regime-switching approach which enables us to deal with shifting behavioral patterns during both high-traffic periods and low-traffic ones. This allowed us to conduct a statistical evaluation and forecast passenger flows at different periods with differing dynamics within a time series describing Sofia citizens' transportation habits.
Public transportation routes and timetables created using data could make public transportation more appealing and favored by citizens while decreasing car trips would improve air quality, noise pollution, and CO2 emissions within our cities.

Limitations and Future Research

This study has several limitations that we acknowledge and suggest avenues for further rese arch. First , this study is limited by the region of the study 's research and results may vary depending on where it is conducted. Performing it elsewhere might yield different findings. Thus, further research could focus on local-specific features and factors influ encing urban travel dynamics, as well as compare existing findings in this study against those in other regions.

Second, the study is constrained by the dataset time period. Another possible study is one utilizing a different dataset in terms of the time period from April 13th until August 31st, 2022. An extended study time period would provide a more general trend in urban mobility and the results could yield further valuable insight into passenger flow dynamics.

Third, this study suggests a framework that works well with a time series of statistically aggregated and anonymized geospatial records of mobile devices that have different frequencies and changing behavior patterns. Although the research framework provides an in-depth view of passenger flow, factors influencing the dynamics of urban mobility may not be completely generalizable from this dataset.

Modeling the share of daily trips taken by each mode of travel can improve the planning of public transport routes and timetables .In a recent paper of S. Lee et al.(2022) city-level modal splitting (the share of daily trips taken by each mode of travel) is explored in terms of city characteristics. Understanding how city features affect modal splitting is crucial to achieving a shift towards low-emission modes of travel and sustainable mobility. The authors note that socio-demographic fac tors have the greatest impact on the modal splits of cities. High population density and high employment rates are also positively associated with low-emission travel modes . People often reconsider owning a private vehicle because of taxes, insurance, and fees on cars, high gasoline prices, low fares for public transportation and taxis, etc. The cost of transportation from point A to point B depends on the type of fuel or energy and its price (diesel, gasoline LPG, NLG, hydrogen, electricity) in comparison to the expenses associated with using public transport. Convenience should also be taken into account.
Overall, these limitations offer researchers who examine passenger flow dynamics with anonymized and statistically aggregated CDR datasets a significant opportunity for further research.

Contact information: lomev@feb.uni-sofia.bg

Contract BG-RRP-2.004-0008 for the financing of project "Sofia University - Marking Momentum for Innovation and Technological Transfer" under pillar 2 "Establishing a network of research higher education institutions in Bulgaria", component "Innovative Bulgaria" from National Recovery and Resilience Plan part of the program to accelerate economic recovery and transformation through science and innovation