Can we classify how people move around a city:

Using travel patterns to predict purpose across the region of Montréal, Canada between September – October 2017.

*This dissertation is submitted in part requirement for the MSc in the Centre for Advanced Spatial Analysis, Bartlett Faculty of the Build Environment, UCL.*

*A colour version of this dissertation is available online.*

Word Count:11,986

Thomas J. Keel – 18110348

TMSSDSAVIS01: MSc Spatial Data Science & Visualisation University College LondonSupervisor: Huanfa Chen

*30th August 2019*

# Abstract

Quantifying the ways in which people travelling and for which purpose, remains an area within broader mobility studies without an extensive investigation. Arguably, this investigation has been hindered by two factors: (1) an absence of relevant data which details how and why people travel in a city and (2) a lack of consistency in the metrics used when quantifying the space-time dynamics of mobility. In recent years, smartphones have provided researchers a platform to study movement within a city at increasingly fine temporal and spatial scales. This, in turn, has fuelled an eruption in sources of volunteered geographic information (VGI) provided by Open Data Initiatives. This study makes uses of one such source: the *2017 MTL Trajet* project in Montreal – a survey examining travel behaviour patterns across Montreal between 18th September 2017 – 18th October 2017. Understanding results from this survey may provide insight into how to better plan essential and non-essential services within a city. This project focuses on discerning spatial and temporal interdependencies using clustering before a comparison of three distinct classification models (SVM, RF and ANN)

**Key Words:** Mobility, Volunteered Geographic Information, Spatio-Temporal Investigation, Networks, Machine Learning, Spatio-Temporal Investigation

# Declaration

I, Thomas Keel, hereby declare that this dissertation is all my own original work and that all sources have been acknowledged. It is 12,000 words in length

Signed:

Date: 28th August 2019

Acknowledgements

Introduction

List of Figures and Tables

List of Acronyms and Tables

Chapter 1. Introduction

Chapter 2. Literature Review

Chapter 3. Methodology

Chapter 4. Results

Chapter 5. Discussion

Chapter 6. Conclusion

Bibliography

Appendix 1. Notification of intent not to apply for ethical approval

Appendix 2. Link to scripts used for the analysis

Table of Contents

[Abstract 1](#_Toc17491468)

[Declaration 2](#_Toc17491469)

[List of Figures and Tables 4](#_Toc17491470)

[List of acronyms and abbreviations 4](#_Toc17491471)

[Acknowledgments 4](#_Toc17491472)

[1. Introduction 4](#_Toc17491473)

[1.2 Motivation 5](#_Toc17491474)

# List of Figures

**Figure 1.1** Screenshot from the MTL Trajet App (left; source: Patterson, 2017) and example of prompt within Itinerum apps (right; source: Patterson *et al.*, 2019)………………………..……….1

**Figure 2.X** Framework of the Itinerum Mobile and Desktop app (Source: Patterson & Fitzsimmons, 2017)…………………………………………………………………………………………………...…………2

**Figure 2.X** Framework for survey apps developed using the Itinerum platform (Source: Patterson & Fitzsimmons, 2017)…………….……………………………………....……………………………………3

**Figure 2.X** Example of GPS trace with location collection priorities within an Itinerum platform app (A); Example of the on screen prompt after an Itinerum platform app stops recording movement (B) (Source: Patterson *et al.*, 2019)……………………………………………………..4

**Figure 3.X** Montreal within Quebec, Canada………………………...………………………..……………………5

**Figure** **3.X** GPS routes from the MTL Trajet plotted within the study area…………...………...……5

**Figure 3.X** Location of Montreal within the study area. ……………………………………………………….5

# List of Tables

**Table 3.1** Description of the key variables from data from the MTL Trajet survey before pre-processing

**Table 3.2** Categories of mode and purpose allowed for trips in the MTL Trajet survey

# List of Acronyms and Abbreviations

**ANN** – Artificial Neural Networks

**CNN** – Convolutional Neural Networks

**DA** – Dissemination Areas

**GPS** – Global Positioning System

**LDA** – Latent Dirichlet Allocation

**MAUP** – Modifiable Areal Unit Problem

**MLP** – Multi-Layer Perceptron

**MTUP** – Modifiable Temporal Unit Problem

**RF** – Random Forest

**SVM** – Support Vector Machine

**VGI** – Volunteered Geographic Information

# Acknowledgments

I would like to thank my family for their thorough support

Secondly, I would like to thank my supervisor Huanfa for their comments and advice and the staff at CASA for the support this year.

Finally, I am extremely grateful for those people I have met throughout this year at CASA, who have

Especially my flat mates James and George.

All dissertations are submitted electronically, but we also require ***two printed and bound copies to be submitted as well***. Double spacing must be used, except for indented quotations, tables, bibliographies and footnotes, which should be single-spaced. The left- side margin should be not less than 40mm (1.5 inches) and other margins not less than 20mm (0.75 inch).

# Introduction

1.1. Overview and Research Questions

The way in which people move travel and the reason for their movement on a large scale, still remains an area with a distinct lack of investigation within broader mobility studies (Yazdizadeh *et al.,* 2019). This somewhat owes to a lack of data available to study this […].

In recent years, smartphones have provided researchers an opportunity to create survey apps that can record large scale geospatial movement patterns (Li et al., 2016). As smartphones exhibit the same mobility behaviour as their carriers (Jahromi *et al.*, 2016). Through mobility surveys we can gain insight into people’s movement behaviour in a city and the preferences of mode of travel they take when travelling for certain activities. Arguably, by studying these behaviours on such a scale, they can use them to inform in city-level decision-making (Attard *et al.*, 2016). For example, if we knew that people have a higher propensity to use public transport when travelling to leisure activities in certain parts of a city, transport authorities could use this information couple bus/metro routes/schedules to locations and opening hours of leisure services [may need better example].

Volunteered Geographic Information (hereafter VGI).

Despite this, there is a gap in knowledge of understanding of movement in most cities, owing to the fact that no research has been initiated there. One exception is in Montreal, where a number of mobile applications have been created in recent years (since 2016) to study how people move across the city based on their smartphones. This report makes use of the most recent available data from one of these studies: The *2017 MTL Trajet* study carried out between 18th September 2017 – 18th October 2017. We use data from this study to assess the following research objectives:

1. Which modes of transports are preferred for which activities carried out throughout the city?
2. Is there a discernible space-time signal of movement and purpose in the city?
3. Can purpose of travel be used as an indicator in models which focus on quantifying movement in cities?
4. Does directional dependence exist in of travel purpose and travel modes?
5. Which form of clustering performs better

# 1.2 Motivation

*What movement is and why studying it is important*

Movement is an interaction between an origin and destination (O/D), whereas transport is a derived demand for people to get to a given location (Golledge & Gärling, 2001; Murray *et al.*, 2012). Studying the purpose of the flow between O/D on a transport network underpins our understanding of the movement behaviour of populations in a city. Indeed, discerning the spatio-temporal profile of the different transport modes and purposes in city can brings us closer to understanding the human phenomena in cities (Kwan & Neutens, 2012). This study aims to deconstruct the spatial, temporal and spatial-temporal aspects of travel […]

Research carried out by Zhang & Cheng (2019) discover expected difference in the profiles of people travelling within London based on their employment status. In general, finding regularity in full-time transport patterns compared with those who are un-employed. While, this information is of use to transport authorities, there is still a lack of investigation into more of the local impacts of transport.

with smartphones enormous potential to collect location data for many purposes, hence ﻿Itinerum platform created (Patterson *et al.*, 2019)

*Current modelling and the state of Big Data in time-space understanding in cities*

Big data measurement and influx has extended to the extent that it is real-time, which gives us a unique opportunity to study geographical phenomena in real time (Goodchild, 2013)

Big geographic data allows for us to not only study the spatial and temporal interactions but also interactions of socio-economic factors [this is what this research aims to do] (Cheng *et al.*, 2017).

Mathematical models being employed without regard of space, often including problems that are inherently tied to spatial considerations (O'Sullivan & Manson, 2015)

*Can purpose of travel be used as an indicator in models which focus on quantifying movement in cities?*

We would have new insights if we new that 90% of travel to hospitals is done by cars for example.

Warrants further investigation

*Smart phones and new ways of studying space-time*

Smartphones capable of environmental geospatial information at an unprecedented level (Li et al., 2016; etc. (i.e. Patterson)).

Also, space-time methods and smart cities

*Why Montreal*

Smart cities [i.e.] the instrumentation of cities which are now providing vast amount of real-time data (from Li *et al.*, 2016). Increase in the amount of geo-tagged/geo-referenced data

Could be important for the allocation of transport planning (ref), resource allocation in a city

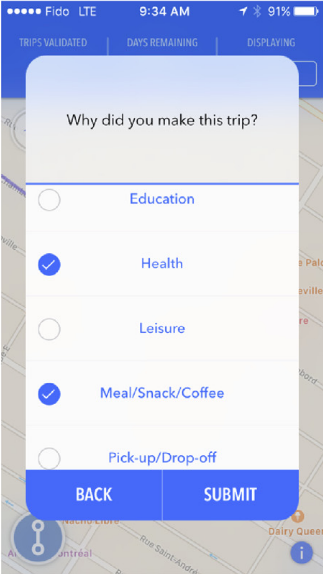
*1.3 Approach*

*MTL Trajet*

This study makes use of data from the *2017 MTL Trajet* survey originally collected by researchers at the Transportation Research for Integrated Planning (TRIP) lab, Concordia University (Patterson & Fitzsimmons, 2017). This survey was created to study travel behaviour and is currently being used to inform city and transport planning decisions in the city of Montreal (MTL Trajet, 2017).

The methodology of this project differs from previous research as it concerns itself with the modelling (through classification) of the movement of people in Montreal based on given purposes and mode. It is hoped that this modelling, in combination with an overarching spatial-temporal investigation, can be used infer something […] about movement within a city.

**Figure 1.1**

** **

**Figure 1.1** Screenshot from the MTL Trajet App (left; source: Patterson, 2017) and example of prompt within Itinerum apps (right; source: Patterson *et al.*, 2019).

Although, this study primarily focusses in on Montréal and this may not be transferred to other cities.

The dataset forming the backbone of this report provides a look into the way that people move.

This study also concerns itself with the modelling through classification on purpose of the movement. It is hoped, in combination with a spatio-temporal investigation, this analysis presented can infer something about movement at a higher scale within a city. Although, it must not be forgotten that this study primarily focusses in on Montréal and this may not be transferred to other cities (Ergodoic and Ecological Fallacy).

This study attempts to break away from its data-driven approach to provide more context

This following examine the related literature and reviews the philosophy of the metrics introduced in 2. After this, in 3 detail is provided on the data . Results in 4 first create a ESTDA before modelling.

First paragraphs:

* Movements can be similarly viewed as spatial interactions between an origin and destination.(Murray et al. 2012)
* ﻿Another reason is that goods are produced in one location, perhaps a factory or farm, then shipped to consumers at other locations. Of course, interaction could also be viewed as trade flows, as considered between towns, cites, regions, states, and countries. (Murray et al. 2012).
* ﻿Sometimes the movement patterns that result are understandable or explainable, like birds migrating south for the winter, but often times they are not obvious (Murray et al. 2012)
* Studying these geographical flows important
* Insight into which activities occur on which days and times (similar to Zhang & Cheng, 2019). -> (lead onto Batty, 2013)
* Characterising and classifying the movement of entities across space or a network is a very interesting area of spatial research that I would like to add to and learn from

Other Ideas:

* Also an examination of where to people drive (and relation to parking spaces) -> this is flawed however as only subset
* Which modes are used for which activities
* big data allows for us to not only study the spatial and temporal interactions but also interactions of socio-economic factors [this is what this research aims to do] (Cheng *et al.*, 2017)
* Wireless portable devices are carried by humans, exhibit the same mobility behaviour of their human carriers and their movements (Jahromi *et al.*, 2016).
* ESSENTIAL: Smart City with interacting networks and GPS signals (Jahromi *et al.*, 2016)
* Also EMBED research questions the study poses

OtheR:

* This somewhat owes to a lack of data and . Primarily the movement of people is of concern to time-space analysis.
* Something about Montreal.
* Real time data gives us ability to study real-time processes within a city Goodchild (2013).
* Smartphones capable of environmental geospatial information at an unprecedented level (Li et al., 2016).
* Understanding mobility through mobile phone has kicked off (Zhao *et al.*, 2019)
* Certain cities have a gap, Montreal
* City planning