INFERRING MODAL SPLIT FROM MOBILE PHONES
Principles, Issues and Policy Recommendations

Norbert Brändle, Austrian Institute of Technology
OVERVIEW DISCUSSION PAPER + THIS TALK

Introduction

Principles of travel mode inference

- Active travel mode inference with smartphone apps
- Passive travel mode inference in mobile phone networks

Issues with travel mode inference from mobile phone data

- Data quality
- Privacy and data protection

Policy recommendations

References
Infer travel mode (foot, bike, car, bus, train, subway, …) from recorded location and sensor data.

**GPS**

- Location:
  - GPS
  - WiFi
  - Cell Data

- Sensor data:
  - Accelerometer
  - Gyroscope
  - Magnetometer

Automatic Travel Diary (Source: AIT Smart Survey)

19/01/2021
EXAMPLE: VELOCITY AND ACCELERATION

- Travel mode inference is a very hard pattern classification problem!
PHENOMEN 1: MULTIPLE INFLUENCE FACTORS

- Person (driving style, gait)
- Travel route (pavement, functional road class…)
- Traffic State
- Type of vehicle (sports car, SUV, …, type of public transport vehicle)
- Tracking device, sensors (range, precision, sampling frequency, GPS availability)
- Position of the device (pocket, bag, in the hand, smart phone holder)
- Interactions with the device (phone calls, …)
- Sitting, standing, walking in public transport
DANGER OF SPURIOUS CORRELATIONS

Per capita consumption of mozzarella cheese correlates with Civil engineering doctorates awarded

Source: www.spuriouscorrelations.com
PHENONEMON 2: SPARSE AND NOISY LOCATION DATA

- Identifying most likely route through transportation network via map matching → costly in terms of memory and computation time

Intermodal journey in Vienna

• Location reports + uncertainty

Provided by operating system, increasingly restrictive ("stalkerware")
EVALUATION OF ACCURACY

• There is currently a lack of comprehensive (standard) reference datasets!

• Example: Sussex-Huawei Locomotion (SHL) Dataset
  • 750 hours of labelled locomotion data
  • from 3 persons

(http://www.shl-dataset.org)
ISSUE: DATA QUALITY

Sussex-Huawei Locomotion (SHL) Challenge 2018, Ubicomp 2018 Singapore
(subset of 272 hours of data from, labels of 60 seconds segments)

- Only 5 submissions did not overfit to training data!
- ... and only data from one person with one carrying position was used for the challenge
POLICY RECOMMENDATIONS (1)

• Fostering the collection of comprehensive reference datasets
  • Data collection campaigns covering all travel modes, combinations of travel modes, vehicle types, …
  • Using 'scripts' instead of daily routine to avoid spurious correlations

• Companies developing smartphone operating systems and controlling the smartphone app marketplace must keep granting access to fine-grained smartphone location data

• Enabling the implementation of reference data exchange frameworks
  • Collection of best practices and guideline, common data formats for intermodal trip data, …
THANK YOU!

Norbert Brändle
Norbert.braendle@ait.ac.at