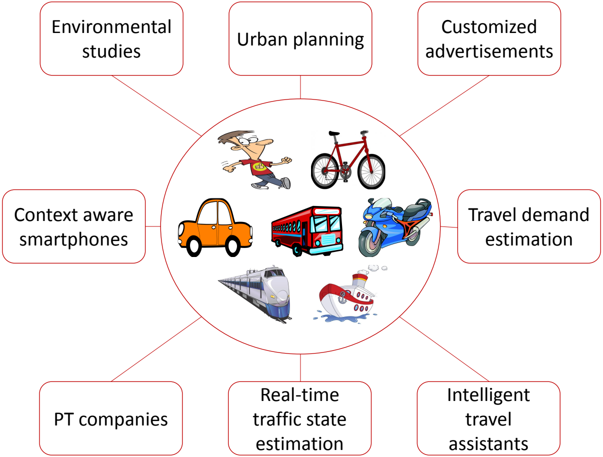
STRC 2017 - 17th Swiss Transport Research Conference, Ascona

Review of transportation mode detection approaches based on smartphone data

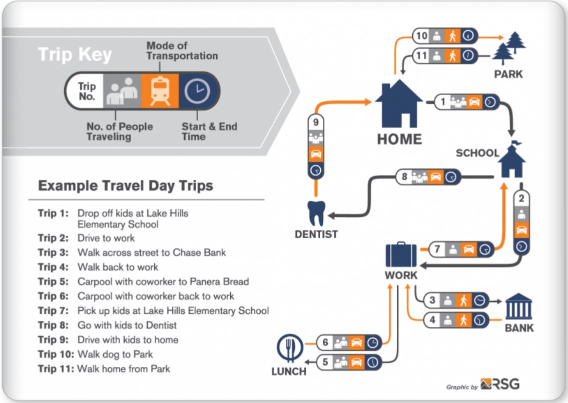
Marija Nikolić, Michel Bierlaire

May 18, 2017

Transportation mode detection (TMD)



Travel surveys



Drawbacks:

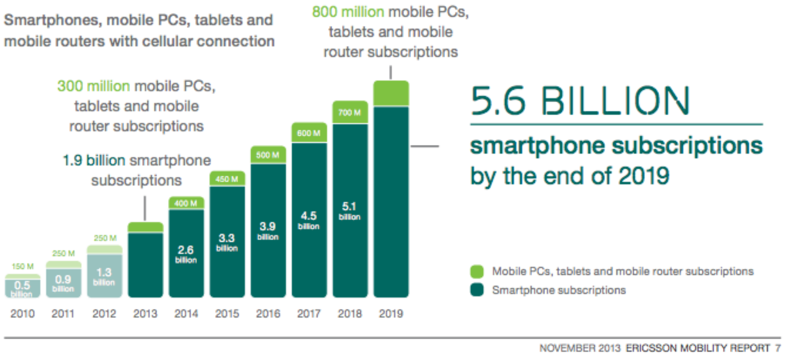
Biased response

No response

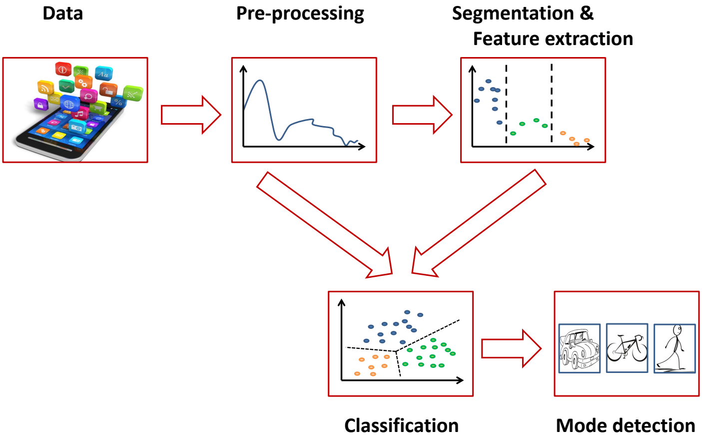
Erroneous reporting

Smartphones: Mobile personal computers

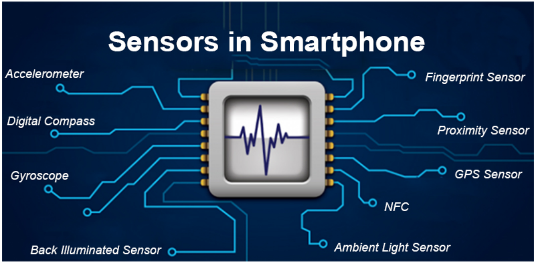
Smartphone penetration



TMD: Procedure



TMD: Sensor data



Motion sensors
Position sensors
Environmental sensors

TMD: Sensor data

Accelerometer

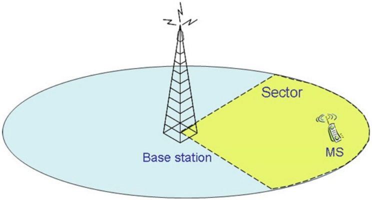
- The acceleration force on all three physical axes
- Independence of any external signal sources
- Low energy consumption

Global Positioning System (GPS)

- The position and velocity information
- Outdoor context
- Reduced precision in dense urban environments
- Modest accuracy (50-80 meters)
- High power consumption

TMD: Sensor data

Cellular network signals: GSM



The fluctuation pattern of cell identifiers and signal strength

- Information on the position, outdoor and indoor contexts
- Precision: 50 - 200 meters, ping-pong effect

Data from mobile phone operators

- Anonymous location measurements, coarse-grained

TMD: Sensor data

WiFi

- Provides wireless connectivity to devices inside a WLAN
- Low positioning accuracy
- The most power-demanding sensor after GPS

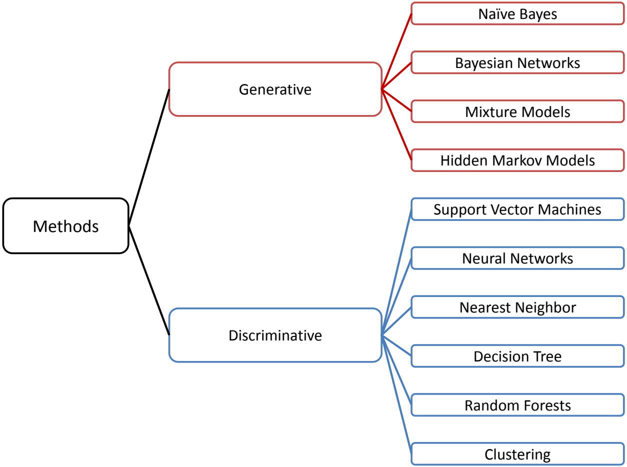
Bluetooth

- Wireless connectivity and short range communication
- Sense devices in their vicinity
- Range: 10 - 100 meters
- Penetration rate: 7 - 11%

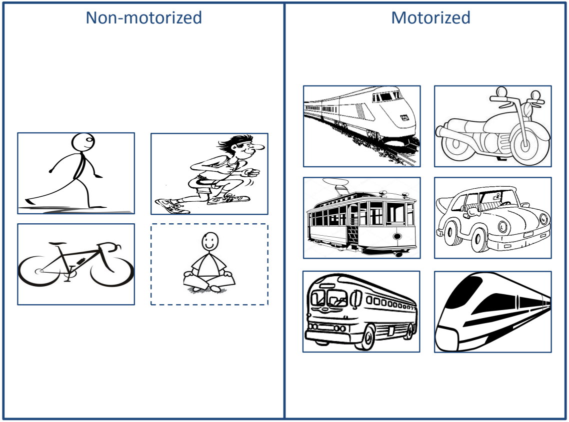
Barometers, thermometers, humidity sensors, cameras...

TMD: External data sources

TMD: Classification algorithms



TMD: Categories



TMD approaches: Comparison

Source	Modes	Smartphone data	External data	Algorithm	Accuracy	
Patterson et al. (2003)	Walking Bus Car	GPS	GIS	Bayes Model	84%	
Muller (2006)	Walking Stationary Car	GSM	/	Artificiel Neural Network **Hidden Markov Model**	Average: 80% Walking: 87% Stationary: 98% Car: 75%	
Sohn et al. (2006)	Walking Stationary Driving	GSM	/	Naïve Bayes Support Vector Machines heuristic-based methods **2-stage boosted Logistic Regression**	Average: 85% Walking: 70.2% Stationary: 95.4% Driving: 84.3%	
Reddy et al. (2008)	Walking Stationary Biking Running Motorized	GPS Accelerometer	/	Naïve Bayes Support Vector Machines Decision Trees k-Nearest Neighbors Continuous Hidden Markov Model **Decision Trees** **and Discrete Hidden Markov Model**	>90%	height
Mun et al. (2008)	Walking Stationary Driving	GSM WiFi	/	Decision Trees	Average: 88% Walking: 90.17% Stationary: 90.26% Driving: 87.83%	
Zheng et al. (2008)	Walking Biking Driving	GPS	/	Graph-based	Average: 76.2% Walking: 89.1% Biking: 66.6% Driving: 86.1%	
Miluzzo et al. (2008)	Sitting Stationary Walking Running	Accelerometer	/	JRIP rule learning	Average: 78.9% Sitting: 68.2% Stationary: 78.4% height Walking: 94.4% Running: 74.5%	

TMD approaches: Comparison

Source	Modes	Smartphone data	External data	Algorithm	Accuracy
Reddy et al. (2010)	Walking Stationary Biking Running Motorized	GPS Accelerometer	/	Naïve Bayes Decision Trees k-Nearest Neighbors Support Vector Machines k-Means Clustering Continuous Hidden Markov Model **2 stage Decision Tree and Discrete Hidden Markov Model**	Average: 93.6% Walking: 96.8% Stationary: 95.6% Biking: 92.8% Running: 91% Motorized: 93.9%
Stenneth et al. (2011)	Walking Bus Car Train Stationary Biking	GPS	GIS	Naïve Bayes Decision Trees Bayesian Network Multilayer Perception **Random Forest**	Average: 93.7% Walking: 96.8% Bus: 88.3% Car: 87.5% Train: 98.4% Stationary: 100% Biking: 88.9%
Xiao et al. (2012)	Mass Rapid Transit Bus Taxi Running	GPS GSM Accelerometer	/	Decision Trees	NA
Montoya et al. (2015)	Walking Biking Bus Train Tram Motorized	GPS WiFI Accelerometer GSM Bluetooth	Road maps Rail maps Public transport schedules Public transport routes	Dynamic Bayesian Network	Average: 75.8% Walking: 91% Biking: 36% Bus: 80% Train and Motorized: 81% Tram: 91%
Chen and Bierlaire (2015)	Walking Biking Car Bus Metro	GPS Bluetooth, Accelerometer	Open Street Map	Probabilistic method	SI>90%
Sonderen (2016)	Walking Running Biking Car	Accelerometer Gyroscope Magnetometer	/	**Decision Tree** Random Forest k-Nearest Neighbors	98%

Comparison: Data sources

- Typically one or two sensors used: accelerometer and GPS
- External data: rarely used (transportation network data)
- Accuracy: higher if more data sources are utilized

Comparison: Classification algorithms

- Generative models: better suited when mobile phones are used only as a sensing system
- Discriminative models: better suited when detection is intended to run on mobile devices directly
    - Decision Trees: satisfactory accuracy while using the least resources

Comparison: Categories & Accuracy

- Predominant: stationary, walking, biking and a unique motorized transport modes
- The best accuracy: walking and stationary modes
- Key challenge: differentiation between motorized classes (bus, car, train, metro)
- External data
    - Added value in detecting various motorized modes
    - Public transportation detection

Comparison: Performance

Generative models: Chen and Bierlaire (2015)

- Probabilistic method: the inference of transport modes and physical paths
- Structural travel model: captures the dynamics of smartphone users
- Sensor measurement models: capture the operation of sensors
- Categories: walking, biking, car, bus and metro
- Smartphone sensors: GPS, Bluetooth, and accelerometer
- External data: transportation network

Comparison: Performance

Discriminative models: Stenneth et al. (2011)

- Random Forests to infer a mode of transportation
- Findings supported by other studies: Abdulazim et al. (2013); Ellis et al. (2014); Shafique and Hato (2015)
- Categories: car, bus, train, walking, biking and stationary
- Smartphone sensors: GPS
- External data: transportation network

Conclusion

- Transportation mode detection based on smartphone data
- The approaches differ in terms of
  - The type and the number of used input data
  - The considered transportation mode categories
  - The algorithm used for the classification task
- Accuracy: higher if more data sources are utilized
- External data: essential for the detection of various motorized modes

Future directions

- Studies with lager samples and over a longer time periods
- Water transportation modes
- Utilization of GSM logs provided by the operators
- Additional data sources

  Barometers, temperature, humidity sensors
  Real time traffic information
  Socio-economic and demographic data
  Mobility and transport census data
  Seasonal data, weather conditions
- Transportation network data: OpenStreetMap
- Public transportation data: opendata.swiss

Thank you

STRC 2017 - 17th Swiss Transport Research Conference, Ascona

Review of transportation mode detection approaches based on smartphone data

Marija Nikolić, Michel Bierlaire

- marija.nikolic@epfl.ch

References I

Abdulazim, T., Abdelgawad, H., Habib, K. and Abdulhai, B. (2013). Using smartphones and sensor technologies to automate collection of travel data, *Transportation Research Record: Journal of the Transportation Research Board* (2383): 44–52.

Chen, J. and Bierlaire, M. (2015). Probabilistic multimodal map matching with rich smartphone data, *Journal of Intelligent Transportation Systems* **19**(2): 134–148.

Ellis, K., Godbole, S., Marshall, S., Lanckriet, G., Staudenmayer, J. and Kerr, J. (2014). Identifying active travel behaviors in challenging environments using gps, accelerometers, and machine learning algorithms, *Public Health* **2**: 39–46.

References II

Miluzzo, E., Lane, N. D., Fodor, K., Peterson, R., Lu, H., Musolesi, M., Eisenman, S. B., Zheng, X. and Campbell, A. T. (2008). Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application, *Proceedings of the 6th ACM conference on Embedded network sensor systems*, ACM, pp. 337–350.

Montoya, D., Abiteboul, S. and Senellart, P. (2015). Hup-me: inferring and reconciling a timeline of user activity from rich smartphone data, *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM, p. 62.

References III

Muller, I. A. H. (2006). Practical activity recognition using gsm data, *Proceedings of the 5th International Semantic Web Conference (ISWC). Athens*, Citeseer, pp. 1–8.

Mun, M., Estrin, D., Burke, J. and Hansen, M. (2008). Parsimonious mobility classification using gsm and wifi traces, *Proceedings of the Fifth Workshop on Embedded Networked Sensors (HotEmNets)*.

Patterson, D. J., Liao, L., Fox, D. and Kautz, H. (2003). Inferring high-level behavior from low-level sensors, *International Conference on Ubiquitous Computing*, Springer, pp. 73–89.

References IV

Reddy, S., Burke, J., Estrin, D., Hansen, M. and Srivastava, M. (2008). Determining transportation mode on mobile phones, *Wearable computers, 2008. ISWC 2008. 12th IEEE International symposium on*, IEEE, pp. 25–28.

Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M. and Srivastava, M. (2010). Using mobile phones to determine transportation modes, *ACM Transactions on Sensor Networks (TOSN)* **6**(2): 13.

Shafique, M. A. and Hato, E. (2015). Use of acceleration data for transportation mode prediction, *Transportation* **42**(1): 163–188.

Sohn, T., Varshavsky, A., LaMarca, A., Chen, M. Y., Choudhury, T., Smith, I., Consolvo, S., Hightower, J., Griswold, W. G. and De Lara, E. (2006). Mobility detection using everyday gsm traces, *International Conference on Ubiquitous Computing*, Springer, pp. 212–224.

Sonderen, T. (2016). Detection of transportation mode solely using smartphones, *Technical report*, University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science.

References VI

Stenneth, L., Wolfson, O., Yu, P. S. and Xu, B. (2011). Transportation mode detection using mobile phones and gis information, *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM, pp. 54–63.

Xiao, Y., Low, D., Bandara, T., Pathak, P., Lim, H. B., Goyal, D., Santos, J., Cottrill, C., Pereira, F., Zegras, C. et al. (2012). Transportation activity analysis using smartphones, *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*, IEEE, pp. 60–61.

References VII

Zheng, Y., Li, Q., Chen, Y., Xie, X. and Ma, W.-Y. (2008). Understanding mobility based on gps data, *Proceedings of the 10th international conference on Ubiquitous computing*, ACM, pp. 312–321.