Transporation mode detection using random forest

Jasna Urbančič
Artificial Intelligence Laboratory,
Jožef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
jasna.urbancic@ijs.si

Veljko Pejović
Faculty of Computer and Information science,
University of Ljubljana
Večna pot 113, 1000 Ljubljana
Slovenia
veljko.pejovic@fri.uni-lj.si

Dunja Mladenić
Artificial Intelligence Laboratory,
Jožef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
dunja.mladenic@ijs.si

ABSTRACT

This paper addresses transportation mode detection for a mobile phone user using machine learning and based on mobile phone sensor data. We describe our approach to data collection, preprocessing and feature extraction. We evaluate our approach using random forest classification with focus on feature selection. We show that with feature selection we can significantly improve classification scores.

1. INTRODUCTION

In the recent years we have witnessed a drastic increase in sensing and computational resources that are built in mobile phones. Most of modern cell phones are equipped with a set of sensors containing triaxial accelerometer, magnetometer, and gyroscope, in addition to having a Global Positioning System (GPS). Smart phone operating system APIs offer activity detection modules that can distinguish between different human activities, for example: being still, walking, running, cycling or driving in a vehicle [2, 3]. However, APIs cannot distinguish between driving in different kind of vehicles, for example driving a car or traveling by bus or by train. Recognizing different kind of transportation, also known as transportation mode detection, is crucial for mobility studies, for routing purposes in urban areas where public transportation is often available, for facilitating the users to move towards more environmentally sustainable forms of transportation [1], or to inspire them to exercise more.

In this paper we discuss the use of random forest in transportation mode detection based on accelerometer signal. We focus on

1. feature extraction, and
2. feature analysis to determine the most meaningful features for this specific problem and the choice of classifier.

Our main contribution is feature analysis, which revealed the impact of each feature to the classification scores.

2. RELATED WORK

While the first attempts to recognize user activity were initiated before smart phones, the real effort in that direction begun with the development of mobile phones having built-in sensors [10], including GPS and accelerometer sensors. There are still some studies that use custom loggers to collect the data [11, 17] or use dedicated devices as well as smart phones [5]. Although GSM triangulation and local area wireless technology (Wi-Fi) can be employed for the purpose of transportation mode detection, their accuracy is relatively low compared to GPS [11], so latest state of the art research is focused on transportation mode detection based on GPS tracks and/or accelerometer data.

Machine learning approaches for transportation mode detection often rely on statistical, time-based, frequency-based, peak-based and segment-based [8] features, however in most cases statistical features and features based in frequency are used [10, 11, 16]. Feature domains are described in Table 1. Statistical, time-based, and spectral attributes are computed on a level of a time frame that usually covers a few seconds, whereas peak-based features are calculated from peaks in acceleration or deceleration. On the other hand, segment-based features are computed on the recordings of the whole trip, which means that they cover much larger scale. Statistical, time-based, and spectral features are able to capture the characteristics of high-frequency motion caused by user's physical movement, vehicle's engine and contact between wheels and surface. Peak-based features capture movement with lower frequencies, such as acceleration and breaking periods, which are essential for distinguishing different motorized modalities. Additionally, segment-based features describe patterns of such acceleration and deceleration periods [8].

Machine learning methods that are most commonly used in accelerometer based modality detection include support vector machines, decision trees and random forests, however naive Bayes, Bayesian networks and neural networks have been used as well [11, 12]. Often these classifiers are used in an ensemble [16]. The majority of algorithms additionally use Adaptive Boosting or Hidden Markov Model to improve the performance of the methods mentioned above [16, 8, 11, 10]. In the last years, deep learning has also been used [6, 14].

Accelerometer-only approach where only statistical features have been used reported 99.8% classification accuracy, however users were instructed to keep the devices fixed position during a trip. Furthermore, only 0.7% of data was labeled as train [11]. State of the art approach to accelerometer-only

Domain	Description
Statistical	These features are include mean, standard deviation, variance, median, minimum, maximum, range, interquartile range, skewness, kurtosis, root mean square.
Time	Time-based features include integral and double integral of signal over time, which corresponds to speed gained and distance traveled during that recording. Other time-based features are for example auto-correlation, zero crossings and mean crossings rate.
Frequency	Frequency-based features include spectral energy, spectral entropy, spectrum peak position, wavelet entropy and wavelet coefficients. These can be computed on whole spectrum or only on specific parts, for example spectral energy bellow 50Hz. Spectrum is usually computed using fast Fourier transform, whereas wavelet is a result of the Wavelet transformation. Entropy measures are based on the information entropy of the spectrum or wavelet [7].
Peak	Peak-based features use horizontal acceleration projection to characterize acceleration and deceleration periods. These features include volume, intensity, length, skewness and kurtosis.
Segment	Segment-based include peak frequency, stationary duration, variance of peak features, and stationary frequency. The latter two are similar to velocity change rate and stopping rate used by [17]. Segment-based features are computed on a larger scale than statistical, time-based or frequency-based features.

Table 1: Feature domains and their descriptions adopted from [8].

transportation mode detection relies on long accelerometer samples. It uses features from all five domains for classification with AdaBoost with decision trees as a weak classifier and achieves 80.1% precision and 82.1% recall [8].

The performance of transportation mode detection systems depends on the effectiveness of handcrafted features designed by the researchers, researcher's experience in the field affects the results. Thus, there have been approaches that use deep learning methods, such as autoencoder or convolutional neural network, to learn the features used for classification. Using a combination of handcrafted and deep features for classification with deep neural network resulted in 74.1% classification accuracy [15].

3. PROPOSED APPROACH

Work flow of the proposed approach is visualized in Figure 1. The first task is data collection. To collect data we use *NextPin* mobile library [4] developed by the Artificial Intelligence Laboratory at Jožef Stefan Institute. Library is embedded into two free mobile applications. The first one is *OPTIMUM Intelligent Mobility* [1]. *OPTIMUM Intelligent Mobility* is a multimodal routing application for three European cities — Birmingham, Ljubljana, and Vienna. The second one is *Mobility patterns* [4]. *Mobility patterns* is essentially a travel journal. Both applications send five second long accelerometer samples every time OS's native activity recognition modules, Google's ActivityRecognition API [2] for Android and Apple's CMMotionActivity API [3], detect that the user is traveling in a vehicle. We use that accelerometer samples for fine-grained classification of motorized means of transportation.

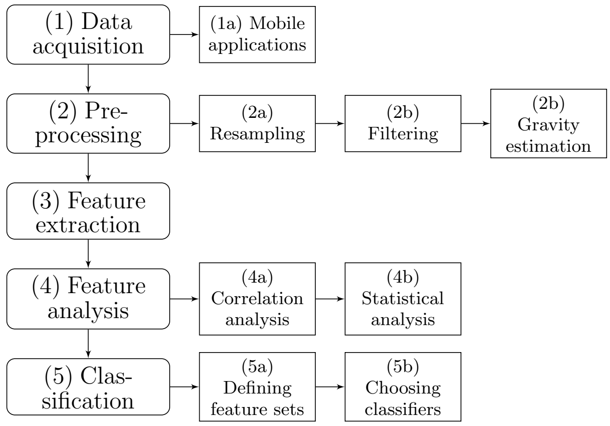

Figure 1: Detailed work flow diagram of the proposed approach. We stacked general, high-level tasks common in other approaches vertically, whereas subtasks specific to our approach are pictured horizontally.

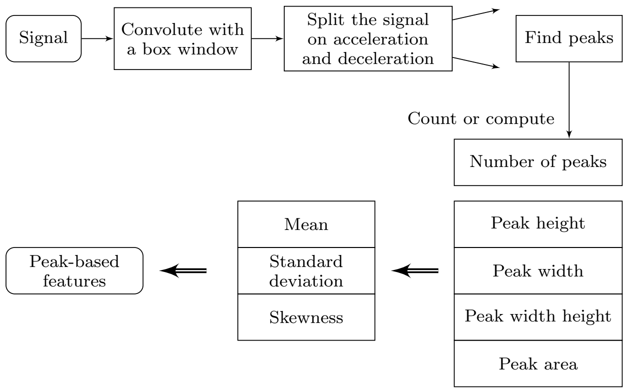

Figure 2: Work flows for extraction of peak-based features.

We collect five second samples of sensor data and resample them to sampling frequency 100 Hz in the preprocessing phase. Resampling ensures us that our samples all contain 500 measurements. The most prominent problem we face in preprocessing concerns the correlation of acceleration measurements with the orientation of the phone in the three dimensional space. Practically this means that gravity is measured together with the dynamic acceleration caused by phone movements. To eliminate gravity from the samples we perform gravity estimation on raw accelerometer signal. We follow an approach proposed by Mizell [9]. Gravity estimation splits the acceleration to static and dynamic component. Static component represents the constant acceleration, caused by the natural force of gravity, whereas dynamic component is a result of user's motion. Furthermore, using this approach we are able to calculate vertical and horizontal components of acceleration.

We use preprocessed signal to extract features for classification. Features are divided into five domains, based on their meaning and method of computation. We have listed the domains in Table 1. We extract features from three domains — statistical, frequency, and peak. We extract statistical features (maximal absolute value, mean, standard deviation, skewness, 5th percentile, and 95th percentile) from dynamic acceleration and its amplitude, horizontal acceleration and

Set	Accele.	Features	Size
D-S	Dynamic	Statistical	54
D-SF	Dynamic	Statistical, Frequency	94
D-SFP	Dynamic	Statistical, Frequency, Peak	172
H-S	Horizontal	Statistical	54
H-SF	Horizontal	Statistical, Frequency	94
H-SFP	Horizontal	Statistical, Frequency, Peak	172
ALL			376

Table 2: Predefined feature sets used for classification.

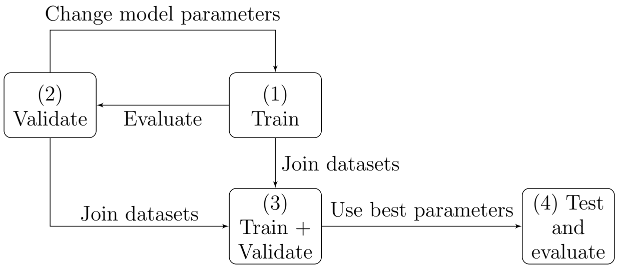

Figure 3: Schema of evaluation scenario.

its amplitude, amplitude of raw acceleration, and amplitude of vertical acceleration. From the same signals we extract frequency-based features using fast Fourier transformation. As frequency-based features we use the power spectrum of the signal aggregated in 5 Hz bins. Pipeline for extraction of peak-based features from dynamic and horizontal in acceleration is pictured in Figure 2. To extract peak-based features we first smooth out the signal with convolution with a box window and split it into moments of acceleration and moments of deceleration. Then we find peaks and compute peak heights, peak widths, peak width heights, and peak areas. As there is usually more than one peak we aggregate these values using mean, standard deviation, and skewness. All together we extract 376 features. We organize features into seven predefined feature sets we use for classification. Feature sets are listed in Table 2.

To evaluate the capabilities and performance of the proposed approach, we divide our dataset in 3 subsets — train, validation, and test set — based on the date the samples were recorded on. By doing so we avoided using in this domain methodologically questionable random assignment of samples collected during the same trip to different subsets. The reason why we did not apply cross-validation is similar. Using samples from the same trip in train and test set would result in significantly higher evaluation scores. For

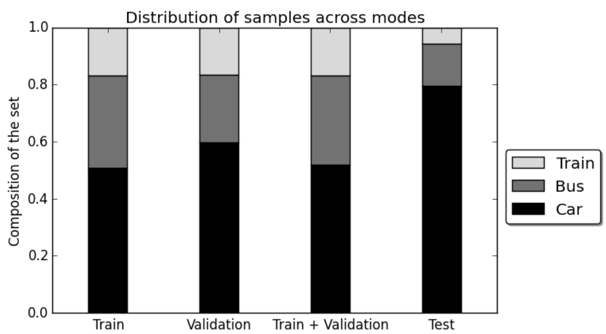

Figure 4: Distribution of modes in train, validation, and test set. We also added joint train and validation set, which we use to train the final model.

Feature set	CA	RE	PR	F1
D-S	0.48	0.41	0.39	0.37
D-SF	0.60	0.41	0.41	0.39
D-SFP	0.46	0.39	0.40	0.35
H-S	0.64	0.40	0.43	0.41
H-SF	0.53	0.39	0.43	0.36
H-SFP	0.50	0.37	0.40	0.34
ALL	0.47	0.35	0.40	0.33

Table 3: Classification metrics for classification with random forest on predefined feature sets.

the training set we use the data from [13], whereas validation and test sets were obtained during Optimum pilot testing in 2018. During validation step we are trying to maximize F1 score as our data set is imbalanced. We visualized the evaluation scenario in Figure 3, while the composition of the sets in pictured in Figure 4.

4. RESULTS

We trained random forest classifier on the predefined feature sets from Table 2. Classification metrics we report on include classification accuracy (CA), recall (RE), precision (PS) and F1 score (F1) Results are listed in Table 3. Table 3 shows that we achieved the highest F1 score of 0.41 using H-S feature set. This feature set consists of statistical features calculated on the horizontal acceleration vector. Classification accuracy in that case is also high, compared to other feature sets. The peak features seems to be the source of noise in the data, as using peak features in combination with the other features sets decreases the performance, for example F1 drops from 0.39 for D-SF to 0.35 for D-SFP.

F1 score and classification for dynamic acceleration increase when we add frequency-based features, whereas these two measures decrease in case of similar action for horizontal acceleration. This offers two possible interpretations. One is that frequency-based features of dynamic acceleration carry more information compared to frequency-based features of horizontal acceleration. The second one is that statistical features of horizontal acceleration are much better than statistical features from dynamic acceleration.

We noticed that smaller feature sets generally perform better than larger so we focused on feature selection. We initially train the model with all features and evaluate it on validation set. Then we remove each feature one by one, train the model, evaluate it on the validation set and compare all F1 scores. We eliminate the feature with the highest F1 score, as this means that when the model was trained without that feature if performed better than when the eliminated feature was included. We repeat this procedure until the feature set consists of one feature. Similarly, we do feature addition — we start with an empty feature set and gradually add features one by one.

Using the described process of forward feature selection and backward feature elimination we selected two feature sets that performed the best — in case of forward selection the best feature set has 10 features, whereas feature set produced with backward elimination has 28 features. Feature set obtained by forward selection mostly contains statistical features, followed by peak-based. Only one frequency-based features appears in that set. Additionally, features are in vast majority extracted from dynamic acceleration. On the other hand feature set obtained by backward elim-

Feature set	CA	RE	PR	F1
Forward selection (10)	0.70	0.50	0.47	0.48
Backward elimination (28)	0.73	0.50	0.48	0.49

Table 4: Classification metrics for classification with the selected features in feature selection.

Forward selection

T \P	Car	Bus	Train
Car	0.78	0.27	0.05
Bus	0.51	0.40	0.09
Train	0.47	0.21	0.32

Backward elimination

T \P	Car	Bus	Train
Car	0.83	0.12	0.05
Bus	0.55	0.35	0.10
Train	0.45	0.23	0.32

Table 5: Confusion matrix for classification with the selected features in feature selection.

ination contains more peak-based features than statistical, again only one frequency-based feature appears. Dynamic acceleration and horizontal acceleration appear in similar proportions. We evaluated the models trained with that feature sets against the test set. Results are listed in Table 4. Both feature sets achieve better F1 scores than any previous feature sets. Confusion matrix in Table 5 reveals what are the differences between these two feature sets. We can see that in case of eliminating features there is less cars missclassified as buses and more buses missclassified as cars. Classification of trains is fairly consistent. For buses and trains the largest part of samples is still missclassified as cars.

5. CONCLUSIONS

We showed that while transportation mode with random forest is possible, careful feature selection is necessary. Using feature selection we were able to improve classification scores for at least 0.04, in some cases even over 0.10. Although classification scores improved, most of non-car samples were still misclassified as cars. We observed that even though peak-based features did not perform as well in predefined feature sets, they appeared consistently among selected features in feature selection. That does not hold for frequency-based feature only one feature appeared among selected features. For the future work we suggest maximization of another classification score as we focused on maximization of F1 score.

6. ACKNOWLEDGMENTS

This work was supported by the Slovenian Research Agency under project *Integration of mobile devices into survey research in social sciences: Development of a comprehensive methodological approach* (J5-8233), and the ICT program of the EC under project *OPTIMUM* (H2020-MG-636160).

7. REFERENCES

[1] Optimum project - European Union's Horizon 2020 research and innovation programme under grant agreement No 636160-2. `http://www.optimumproject.eu/`, 2017. [Online; accessed 4-November-2017].

[2] `ActivityRecognition`. `https://developers.google.com/android/reference/com/google/android/gms/location/ActivityRecognition`, 2018. [Online; accessed 31-August-2018].

[3] `CMMotionActivity`. `https://developer.apple.com/library/ios/documentation/CoreMotion/Reference/CMMotionActivity_class/index.html#//apple_ref/occ/cl/CMMotionActivity`, 2018. [Online; accessed 31-August-2018].

[4] L. Bradeško, Z. Herga, M. Senožetnik, T. Šubic, and J. Urbančič. Optimum project: Geospatial data analysis for sustainable mobility. In *24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining Project Showcase Track*. ACM, 2018. `http://www.kdd.org/kdd2018/files/project-showcase/KDD18_paper_1797.pdf`.

[5] K.-Y. Chen, R. C. Shah, J. Huang, and L. Nachman. Mago: Mode of transport inference using the hall-effect magnetic sensor and accelerometer. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(2):8, 2017.

[6] S.-H. Fang, Y.-X. Fei, Z. Xu, and Y. Tsao. Learning transportation modes from smartphone sensors based on deep neural network. *IEEE Sensors Journal*, 17(18):6111–6118, 2017.

[7] D. Figo, P. C. Diniz, D. R. Ferreira, and J. M. Cardoso. Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing*, 14(7):645–662, 2010.

[8] S. Hemminki, P. Nurmi, and S. Tarkoma. Accelerometer-based transportation mode detection on smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, page 13. ACM, 2013.

[9] D. Mizell. Using gravity to estimate accelerometer orientation. In *Proc. 7th IEEE Int. Symposium on Wearable Computers (ISWC 2003)*, page 252. Citeseer, 2003.

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

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

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

[13] J. Urbančič, L. Bradeško, and M. Senožetnik. Near real-time transportation mode detection based on accelerometer readings. In *Information Society, Data Mining and Data Warehouses SiKDD*, 2016.

[14] T. H. Vu, L. Dung, and J.-C. Wang. Transportation mode detection on mobile devices using recurrent nets. In *Proceedings of the 2016 ACM on Multimedia Conference*, pages 392–396. ACM, 2016.

[15] H. Wang, G. Liu, J. Duan, and L. Zhang. Detecting transportation modes using deep neural network. *IEICE TRANSACTIONS on Information and Systems*, 100(5):1132–1135, 2017.

[16] P. Widhalm, P. Nitsche, and N. Brändie. Transport mode detection with realistic smartphone sensor data. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 573–576. IEEE, 2012.

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