Determining Transportation Mode On Mobile Phones

Sasank Reddy, Jeff Burke, Deborah Estrin, Mark Hansen, Mani Srivastava
University of California, Los Angeles (UCLA)
{sasank, jburke, mbs}@ucla.edu, destrin@cs.ucla.edu, cocteau@stat.ucla.edu

Abstract

As mobile phones advance in functionality and capability, they are increasingly being used as instruments for personal monitoring. Applications are being developed that take advantage of the sensing capabilities of mobile phones - many have accelerometers, location capabilities, imagers, and microphones - to infer contextual information. We focus on one type of context, the transportation mode of an individual, with the goal of creating a convenient (no requirement to place sensors externally or have specific position/orientation settings) classification system that uses a mobile phone with a GPS receiver and an accelerometer sensor to determine if an individual is stationary, walking, running, biking, or in motorized transport. The target application for this transportation mode inference involves assessing the hazard exposure and environmental impact of an individual's travel patterns. Our prototype classification system consisting of a decision tree followed by a first-order Hidden Markov Model achieves the application requirement of having accuracy level greater than 90% when testing with our dataset consisting of twenty hours of data collected across six individuals.

1. Introduction

Mobile phones are truly ubiquitous. They are one of the few devices, which have computation, sensing, and communication capabilities, that are carried by people throughout the day. These devices are being integrated with sensors that capture location and measure acceleration. We are interested in using mobile phones to determine the transportation mode of an individual - whether the user is stationary, walking, running, biking or in motorized transport[1]. Demand for this fine-grained inference exists: our Personal Environmental Impact Report (PEIR) project uses transportation mode tagged location traces as input into models of hazard exposure and environmental impact [1].

This paper outlines our work in creating a transportation mode classification system that runs on a mobile phone equipped with a GPS receiver and a 3-axis accelerometer. Our initial results, using a dataset of twenty hours of data from six individuals, shows promise in creating a convenient (single sensing unit, can be on or inside of clothes, no orientation/position requirement), reliable (greater than 90% accurate) classification system for these states. The system classifies one second of data by using the GPS receiver speed value and energy, variance, and sum of FFT coefficients between 1-5Hz from the accelerometer and employs a decision tree followed by a first-order discrete Hidden Markov Model. Our system could work on any platform that contains a GPS receiver and a 3-axis accelerometer but our design is informed by plausible sensors and position/orientation requirements of mobile phones.

2. Related Work

Existing work related to transportation mode classification has focused on using location information coupled with external data (e.g. road/transportation infrastructure, user patterns, GSM cell tower/WiFi access point density) [2, 3, 4, 5, 6], using accelerometer placed in one or more (up to twelve) positions on the body [7, 8, 9, 10], or using a single sensing unit but with multiple modalities (accelerometer, audio, barometric pressure) [11, 12]. Table 1 shows recent work that has been specifically implemented on mobile phones.

Existing systems do not meet all of our design goals. They are either not convenient (require specific position/orientation of sensors, have to be worn externally), rely on additional information (learned user patterns, land use data), or use infrastructure that is not always readily available (dense GSM cell towers/WiFi access points). Regardless, we use this existing work to help determine what types of sensing modalities, features, and techniques could be useful for our system.

[1] We do not distinguish between motor vehicles (motorcycle, car, bus) here, but are exploring map-matching for this purpose.

	Classes	Sensor	Size	Accuracy
[6]	Still, Walk, Motorized	GSM	1 User, 45 Mins.	82%
[3]	Still, Walk, Motorized	GSM	3 Users, 323 Hours	85%
[5]	Still, Walk, Motorized	GSM, WiFi	2 Users, 13 Hours	83%
[2]	Walk, Motorized	GPS, GIS	1 User, 60 Days	84%

Table 1. Related Work Implemented on Mobile Phones

3. Design Goals

The design goals for our transportation mode classifier included user convenience and meeting the accuracy demands for PEIR. To enable user convenience, we required our system to have the following properties: a.) contained in one sensing unit, b.) flexible in terms of position, orientation, c.) wearable externally or contained in clothing, d.) able to work for a variety of users, and e.) effective with sensors that exist on mobile phones.

The PEIR application derives statistics based on the transportation mode inferences made for an individual, and the allowable noise from the transportation mode classification cannot exceed 10% (accuracy of the classifier has to be greater than 90%). By analyzing several weeks of transportation mode activity of members in PEIR, we have found that higher error rates compromise an individual's ability to make choices about their daily transportation habits; in effect adding noise to the impact/exposure estimates that is on par with "natural" variations that they may want to study (changes in speed or the selection of alternative routes).

4. Approach

4.1. Hardware Platform and Sensors

Since our activities are kinetic based and our system needs to operate when worn in a variety of ways, we employ an accelerometer and a GPS receiver as our sensors. Specifically, we use the Nokia N95 for our data collection, which contains a three axis accelerometer with a sensitivity of +-2G and bandwidth of 35 Hz and a built-in GPS receiver that can sample at 1 Hz [13].

Accelerometer and GPS information is complementary. In situations where the accelerometer output is similar, the speed is typically different and vice versa. When we employed just one of these sensors for classification, we obtain a drop in accuracy of 10% compared to using both. Thus, both modalities are needed.

4.2. Feature Selection

We take a window of 1 second, with an overlap of 0.5 seconds, as our period of classification. Smaller window sizes causes classification accuracy to suffer due to certain features (accelerometer frequencies) not being effective and larger window sizes introduces noise since multiple activities could exist.

Since we assume a random and possibly changing orientation, we take the magnitude of the force vector by combining the measurements from the all 3 axis as the basis for our accelerometer features. We evaluated various features including the mean, variance, energy, and several banks of filters (ranging from 0.5-10Hz with different divisions) based on magnitude of the accelerometer along with the speed of the GPS receiver [7, 9]. We concentrated on frequencies between 0.5-10Hz for the accelerometer since previous work in biomechanics indicates this range is appropriate to detect pedestrian motion [14]. In terms of speed, we use the value obtained from the GPS receiver when possible, which is more accurate than calculating speed from location points.

We selected variance, energy, sum of FFT coefficients between 1-5 Hz from the accelerometer and the speed from the GPS receiver as our feature set using correlation based feature selection (CFS). CFS employs a correlation measure to evaluate the goodness of feature subsets and is based on the idea that a good feature subset contains features highly correlated to a class, yet uncorrelated with each other [15]. Alternative approaches exist - a weighted scheme based on cost of obtaining features or a tiered setup where one feature's value affects which other features to obtain could be implemented. We leave this as future work (Section 6).

4.3. Data Collection

The data set used for training/testing our classifiers was obtained by asking six individuals, three male and three female between the ages of 20-28, to gather eight minutes of data while outside for each of the five transportation modes. The volunteers performed the activities with five phones attached simultaneously - positioned on the waist, chest, hand, pocket, and in a bag. Orientation and attachment procedures were decided by the individual. Instructions were given on the sequence of activities to perform, and an external entity captured the ground truth labels. The total amount of data collected across all six individuals was twenty hours. Note that we required individuals to traverse both typical street and highway roads, while minimizing idling, for the motorized transport case.

5. Results

5.1. Comparing Classifiers

To determine which classification system is the most accurate, we compared: a.) instance classifiers such as K-Nearest Neighbor (kNN), Naive Bayes (NB), C4.5 Decision Trees (DT), and Support Vector Machines (SVM), b.) continuous Hidden Markov Model (CHMM), and c.) two-stage system involving a DT and a discrete Hidden Markov Model (DHMM).

To train/test our classifiers, we used the data set described in Section 4.3 and employed 10-fold cross validation where each fold contained equal amounts of continuous segments from each activity. The Weka toolkit and custom HMM implementations were used for evaluation purposes [16], and the final classification system is implemented directly on the mobile phone using Python. The classification system on the phone uses on average 430 mW of power which translates to roughly 8.15 hours of outdoor operation on the phone's 950 mAh battery. We discuss methods to improve the energy efficiency of our system in Section 6.

5.1.1. Parameters and Specification

In the kNN model, the 12 nearest neighbors were chosen based on comparing the accuracy of various "k" values using cross validation. The SVM used a linear basis function for the hyperplanes. A DT consisting of 21 nodes and a depth of 5 levels was employed by employing reduced error-pruning to avoid over-fitting [16].

The CHMM has five hidden states corresponding to the transportation modes and the output symbols are the accelerometer and GPS receiver features modeled as independent Gaussian distributions. The two-stage classifier is a DT followed up by a DHMM where the DT directly uses the raw features and the DHMM is trained by the class posterior probabilities of the DT. Thus, the DHMM output symbols are the DT classifications and the hidden states are the transportation modes. We set the state transition probabilities for the HMMs to reflect which transitions are rare and which would be more common. A similar two-stage approach was used by [12], but we explore a smaller features space (4 vs 50) and DTs instead of decision stumps, due to higher classification accuracy, for the initial stage.

5.1.2. Classification Accuracy

The accuracy (percentage of correctly classified states) of each type of classifier is shown in Table 2. The two-stage system consisting of the DT and the DHMM was the most accurate (greater than 98%) which is plausible since the DT is tuned to differentiate between the boundaries of transportation modes and the DHMM eliminates noise based on temporal knowledge of the previous transportation mode and the likelihood of transitioning into the next mode.

	Still	Walk	Run	Bike	Motor	All
NB	96.0	87.1	98.4	61.2	93.6	87.2
DT	98.2	96.2	98.6	91.2	94.3	95.7
kNN	97.5	95.2	98.4	91.0	91.2	94.7
SVM	97.8	95.6	98.2	86.9	88.4	93.4
CHMM	96.2	96.1	98.4	89.4	91.7	94.4
DT-DHMM	**98.2**	**99.5**	**99.4**	**98.3**	**98.7**	**98.8**

Table 2. Accuracy Results for Classifiers

5.2. Device Position Variation

Given our goal of user convenience, we investigated how phone placement affects transportation mode recognition accuracy [7, 11]. For testing purposes, we created a general DT trained on data from all five positions (arm, bag, chest, hand, and pocket) and then individual DTs that trained on data from specific positions. The results, shown in Table 3, indicate that the general DT is on par with position specific ones (average decrease of 1.2% in accuracy).

All Positions	95.7
Arm	97.1
Bag	96.9
Chest	96.1
Hand	96.2
Pocket	98.0

Table 3. Phone Position and Accuracy

Also, our analysis shows that user input to filter possible positions can help improve accuracy. For instance, a classifier based on the phone being in a bag, hand, or pocket resulted in 96.2% accuracy, chest and hand in 95.9%, and bag and hand in 96.0%.

5.3. User Variation

Another goal related to user convenience is whether a generic classifier could be built that is effective for new users without additional training [7, 11]. To test the feasibility of such a system, we perform "leave one user out" testing, where we train a DT classifier with all but one user (five out of six) and test with the user not in the training set. In this test, we achieved an average accuracy of 93.2% and a minimum accuracy of 87.7%. Also, we performed a test where we add more individuals (one to six) into a training set while testing on data from all individuals [11]. The results show that performance increases as we introduce more individuals into the training set with accuracy gain stabilizing above 95% after four users.

6. Discussion

The results derived from our user base of six individuals are very promising - we have shown that our classification system is accurate regardless of position/orientation of sensors and that a generic classifier is feasible. But our findings are preliminary and for our results to be more generalizable tests need to performed based on a larger more varied user base. We plan to perform such a data collection as future work.

Reviewing our classification technique, there is opportunity to further tune model parameters. For instance, we chose a frequency range of the 1-5Hz for the FFT of the accelerometer based on optimizing to distinguish between all classes, but an alternative is to use the speed feature to pick the appropriate frequency range.

Another area of further work comes in making our classification method more energy efficient. Currently, our system classifies every second but this might not be totally necessary. [17] suggests that we could use selective sampling techniques, such as ones based on entropy, and still achieve high accuracy. Also, we could consider the cost of capturing and processing of features to control the tradeoff between energy and accuracy.

Finally, we want to explore whether a generic classifier is the best approach to deal with user variation. We would like to consider alternatives such as creating several classifiers that are tuned on user-specified parameters (e.g. likely transportation modes, physical/demographic attributes) or employing a short user-specific training phase. These methods could lead to better performance but have disadvantages as well such as longer startup time and increased user involvement.

7. Conclusion

We created a transportation mode classification system, employing a DT followed by a DHMM, that distinguishes between being stationary, walking, running, biking, and in motorized travel using a mobile phone equipped with a GPS receiver and an accelerometer. We have shown that such a system can be convenient for a user by not having strict position/orientation requirements and allowing the device to be worn outside or inside of clothes while still meeting application demands, accuracy greater than 90%, based on a dataset of twenty hours of data from six users. Our work is just a first exploration - further testing is needed to validate our results and there exists opportunities for expansion.

References

[1] E. Agapie et. al. Seeing Our Signals: Combining Location Traces and Web Based Models for Personal Discovery. *HotMobile*, 2008.

[2] L. Liao, et. al. Learning and inferring transportation routines. *AI*, pages 311–331, 2007.

[3] T. Sohn et. al. Mobility Detection Using Everyday GSM Traces. *Ubiquitous Computing*, pages 212–224, 2006.

[4] E. Welbourne, et. al. Mobile context inference using low-cost sensors. *LOCA*, 2005.

[5] M. Mun, et. al. Parsimonious Mobility Classification using GSM and WiFi Traces. *Hot-EmNets*, 2008.

[6] I.A.H. Muller. Practical Activity Recognition using GSM Data. *ISWC*, 2006.

[7] L. Bao and S.S. Intille. Activity Recognition from User-Annotated Acceleration Data. *Pervasive*, 2004.

[8] S.E. Crouter et. al. Validity of 10 Pedometers for Measuring Steps, Distance, and Energy Cost. *MSSE*, 2003.

[9] N. Kern et. al. Multi-sensor Activity Context Detection for Wearable Computing. *EUSAI*, 2003.

[10] C. Randell and H. Muller. Context awareness by analysing accelerometer data. *ISWC*, 2000.

[11] J. Lester et. al. A Practical Approach to Recognizing Physical Activities. *Pervasive*, 2006.

[12] J. Lester et. al. A hybrid discriminative/generative approach for modeling human activities. *IJCAI*, 2005.

[13] Nokia. Nokia n95. http://www.nseries.com, 2007.

[14] D.A. Winter. *Biomechanics and motor control of human movement*. Wiley New York, 1990.

[15] M.A. Hall. *Correlation-based Feature Selection for Machine Learning*. PhD thesis, Univ. of Waikato, 1999.

[16] I.H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. MKP, 2005.

[17] A. Krause, et. al. Trading off Prediction Accuracy and Power Consumption for Context-Aware Wearable Computing. *ISWC*, 2005.