Mobility Detection Using Everyday GSM Traces

Timothy Sohn[1], Alex Varshavsky[2], Anthony LaMarca[3], Mike Y. Chen[3], Tanzeem Choudhury[3], Ian Smith[3], Sunny Consolvo[3], Jeffrey Hightower[3], William G. Griswold[1], and Eyal de Lara[2]

[1] Computer Science and Engineering
University of California, San Diego, La Jolla, CA, USA
{tsohn, wgg}@cs.ucsd.edu
[2] Computer Science, University of Toronto, Toronto, Canada
{walex, delara}@cs.toronto.edu
[3] Intel Research, Seattle, WA, USA
{anthony.lamarca, mike.y.chen, tanzeem.choudhury, ian.e.smith, sunny.consolvo, jeffrey.r.hightower}@intel.com

Abstract. Recognition of everyday physical activities is difficult due to the challenges of building informative, yet unobtrusive sensors. The most widely deployed and used mobile computing device today is the mobile phone, which presents an obvious candidate for recognizing activities. This paper explores how coarse-grained GSM data from mobile phones can be used to recognize high-level properties of user mobility, and daily step count. We demonstrate that even without knowledge of observed cell tower locations, we can recognize mobility modes that are useful for several application domains. Our mobility detection system was evaluated with GSM traces from the everyday lives of three data collectors over a period of one month, yielding an overall average accuracy of 85%, and a daily step count number that reasonably approximates the numbers determined by several commercial pedometers.

1 Introduction

This paper introduces a technique for detecting a user's coarse-grained mobility using commodity cell phones. Pervasive computing applications have long made use of technologies for inferring a user's physical activities. Both coarse and fine-grained location systems have been used to perform location-driven activity inference [14, 34]. Smart spaces containing cameras, RFID tags, and the like, have been used to detect fine-grained user activities [3, 15, 25]. Unfortunately, the cost, complexity and maintenance overhead of such activity inference systems have hampered their mainstream adoption. Recent work has attempted to address some of these issues. An example is the belt-worn cluster of sensors developed by Lester et al. that can identify several physical activities including detecting subtle distinctions such as walking on level ground versus up stairs [19]. However, challenges in form factor and power usage still remain.

Fortunately, many applications do not require the detail and accuracy of the systems cited above. As an example, consider the domain of eldercare as depicted by Computer-Supported Coordinated Care (CSCC) [7]. CSCC describes the network of

P. Dourish and A. Friday (Eds.): Ubicomp 2006, LNCS 4206, pp. 212 – 224, 2006.

people who help an elder *age in place* and seeks to improve the quality of her care while reducing the burden of providing care on the members of her network. Many of the elder's activities that are meaningful for her network members to know about involve high-level information about the elder such as whether or not she was up and about today, or if she had a sedentary day around the house.

The most widely deployed and used mobile computing device today is the mobile phone, which presents an obvious opportunity for high-level activity recognition such as that needed by CSCC applications. This paper investigates how without GPS, a commodity GSM phone could infer such high-level information without placing the types of additional burdens on the user that are typical of more heavyweight systems. Previous research has used GPS to detect the modes of transportation for an individual [23]. However, GPS positioning is available as little as 5% of a typical person's day [18], providing much lower coverage than we require. In contrast, cellular coverage is available throughout most, if not all, of a person's day and does not require line of sight to work [18]. Therefore, using a GSM sensor to detect high-level activities allows the sensing system to always be available, and allows users to continue to carry their mobile phones in their pockets, bags, etc.

The contribution of this paper is that, with unmodified GSM mobile phones and without relying on users to modify their behavior, we can recognize several high-level activities. Using statistical classification and boosting techniques, we successfully distinguished if a person is walking, driving, or remaining at one place with 85% accuracy. Additionally, we were able to build a GSM-based step count predictor that provides a reasonable approximation of the user's daily step count compared to several commercial pedometers. Our methods were tested with real-world data from three data collectors using the two major GSM networks in the United States (T-Mobile and Cingular). The data collectors gathered GSM network trace data over a period of one month, logging a total of 249 walking events and 171 driving events. Our methods show that GSM-based sensing from commodity devices may provide enough activity information for some applications, without the overhead of requiring additional sensors.

The remainder of this paper is organized as follows. Section 2 describes our algorithms to infer mobile activities and daily step counts. Our data collection, metrics, and evaluation results are presented in Section 3. Section 4 describes several application domains that could benefit from our mobility detection technique. Section 5 outlines related work, and we conclude in Section 6.

2 Mobility Detection with GSM

In this section, we offer a brief overview of the Global System for Mobile Communication (GSM) and describe algorithms that use traces of GSM signals to infer modes of mobility and to estimate daily step-count.

2.1 Global System for Mobile Communication (GSM)

GSM is the most widespread cellular telephony standard in the world, with deployments in more than 200 countries. As of September 2005, the GSM family of technologies has 1.5 billion subscribers and 78% of the world market [1]. A GSM base station is typically equipped with a number of directional antennas that define sectors of coverage, or cells. Each cell is allocated a number of physical channels based on the expected traffic load and the operator's requirements. Typically, the channels are allocated in a way that both increases coverage and reduces interference between cells.

We wrote a custom application for the Audiovox SMT 5600 mobile phone to measure and record the surrounding GSM radio environment. Each reading includes signal strength values, cell IDs and channel numbers of up to seven nearby cell towers. In addition, we extract channel numbers and associated signal strength values of up to 15 additional channels. Cell IDs are uniquely identified by the combination of Mobile Country Code (MCC), Mobile Network Code (MNC), Location Area Code (LAC), and cell id. Although other cell towers may be present in the area, our application only sees those associated with the phone's SIM card provider. We sampled our GSM radio environment with the mobile phone at a rate of one sample per second (1 Hz).

2.2 Inferring User Mobility Modes

Our method for detecting user mobility is based on the same principle as fingerprint-based location systems [5, 22]: namely that the radio signals observed from fixed sources are consistent in time, but variable in space. Thus, given a series of GSM observations with a stable set of towers and signal strengths, we conclude that the phone is not moving. Similarly, we interpret changes in the set of nearby towers and signal strengths as indicative of motion.

We conducted a simple controlled experiment to determine how the radio environment changes as a result of various movement activities. Fig. 1 shows the average Euclidean distance values between consecutive GSM measurements, as the data collectors stood still, walked and drove at different speeds. Conceptually, Euclidean distance captures the similarity between GSM measurements. The smaller the Euclidean distance between two measurements, the more similar these measurements are. For example, if measurement A has 3 cells/channels with signal strengths $\{R_1^A, R_2^A, R_3^A\}$ and measurement B has the same 3 cells/channels with signal strengths $\{R_1^B, R_2^B, R_3^B\}$, the Euclidean distance between measurements A and B will be calculated as:

$$\sqrt{(R_1^A - R_1^B)^2 + (R_2^A - R_2^B)^2 + (R_3^A - R_3^B)^2}$$

If a particular cell/channel is not present in one of the measurements, we substitute its signal strength with the minimal signal strength found in this measurement.

Figure 1 shows that the Euclidean distance between consecutive measurements is proportional with the speed of movement. During stationary periods, the distance values stay relatively small (< 5). The slow and fast walking periods show a distinct difference from the stationary period. The driving traces show the most rapid changes in the radio environment, greater than either walking or stationary. Fast walking and

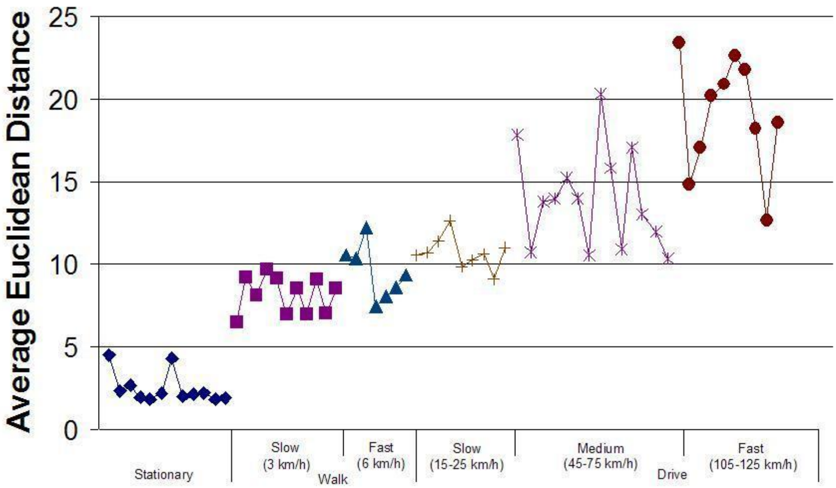

Fig. 1. Average Euclidean distance between consecutive measurements during a stationary period, slow/fast walking periods and slow/medium/fast driving periods

slow driving sometimes overlap in their range of Euclidean distance values, which may result in false recognition between the two states. For a given speed, the Euclidean distance values are not constant because changes in signal strengths are both a function of speed as well as the physical environment, such as buildings, people, or vehicles.

Based on these findings we extracted a set of seven different features to use in classifying a set of GSM measurements as either *stationary*, *walking*, or *driving*. Three features compare two consecutive measurements in time, while the other four features use a sliding window of measurements. We used window sizes of 10, 60, and 300 seconds. Our seven features are:

1. Euclidean distance between two consecutive measurements
2. Spearman rank correlation coefficient [36] between two consecutive measurements. (This number represents how closely the signal strengths from common cell towers were ranked. A more similar ranking indicates less movement.)
3. The number of common cell towers between two consecutive measurements.
4. Mean Euclidean distance over a window of measurements where the values are calculated between consecutive measurements and then averaged together.
5. Variance in Euclidean distance values over a window of measurements where the values are calculated between consecutive measurements.
6. The variance in signal strengths for each tower seen within a given window. (The variance values for each tower are averaged together to produce a single number representing the signal strength "spread" over the entire window.)
7. Euclidean distance value between the first and last measurement of a window.

We used these features to train a two-stage classification scheme. The first stage classified an instance as stationary or not. If the instance was classified as not stationary, a second classifier would determine if the instance was walking or driving. Both classifiers were trained using a boosted logistic regression technique [13] using decision stumps—a single node decision tree. All algorithms were provided by the Weka machine learning toolkit [37]. We chose to use boosting because it has been shown to work well in a variety of classification tasks [24, 27]. In our own experiments we compared boosted logistics regression with naïve Bayes, Support Vector Machines, AdaBoost [12], MultiBoost [35], and some heuristic-based methods; the boosted logistics classifier provided the best recognition rates. We also compared the two-stage classification approach to a single multi-class approach and found that the two-stage classifiers resulted in better accuracy. This is consistent with the findings of Viola and Jones [32], which showed that cascades of classifiers can achieve better recognition rates than single multi-class approaches in face detection tasks. The other advantage of the boosted logistic regression technique using decision stumps is that after the boosting process we have a ranking of features based on how useful they are during classification. Thus the system can be used to select features as well as learn classifiers simultaneously [19, 33]. Furthermore, using only a small subset of the most relevant features can provide computational savings, which is especially important when running inference on a mobile phone.

2.3 Estimating a User's Daily Step-Count

A nice feature of a mobile phone being able to determine periods when a user is walking is that it can be used to approximate how much a user walks, similar to the information provided by a pedometer. Pedometers are currently popular and are used worldwide as a tool to help people track the number of steps they take each day. The benefits of walking and the use of pedometers have been widely promoted by the healthcare community, and a popular suggestion is for people to walk at least 10,000 steps/day [32].

To provide a reasonable measurement of steps taken (or "step count"), a pedometer is clipped to the user's waistband, above the thigh's midline. This restriction may be problematic, as some users do not like the look of the pedometer, or may not have a place to clip it, for example, if the user is wearing a dress. The mobile phone does not have such a restriction, as it can be anywhere with the user, including in her bag. We do not expect a person to always have their phone on her, such as when she is at home. However, being able to provide pedometer-like functionality when outside the home can be useful to give a high-level report of a person's mobile activity for the day.

The GSM-based mobility recognition from the previous section allows us to add a pedometer-like capability to mobile phones. By totaling the number of walking periods and multiplying by an appropriate step rate, we can estimate the user's daily step count. Although this method of calculating step count may seem crude and prone to error, we show in Section 3.3 that our GSM-based step count estimates can approximate that of several commercially available pedometers.

3 Experimental Evaluation

In this section, we evaluate our mobility mode detection and step-count algorithms using data collected from three people to demonstrate the feasibility of using GSM traces to recognize high-level activities. We first describe our metrics and performance for mobility detection. We then evaluate our ability to estimate a user's daily step count.

3.1 Data Trace Collection

Three members of our research team collected GSM network traces as they went about their daily lives for one month. Each data collector carried a commodity GSM phone, the Audiovox SMT 5600, running our software for recording readings from nearby cell towers. Two of the data collectors used Cingular, and one used T-Mobile, spanning the two major GSM network providers in the U.S.

Data collectors recorded their mobility activities using a custom diary application running on the phone that allowed them to indicate whether they were walking, driving or in one place. Each collector also carried a paper notebook where he could record any event that he forgot to indicate on the mobile phone. These paper logs were later transcribed and merged with the digital log for a complete self-reported ground truth. There were a total of 53 corrections (7% of all events) from the paper logs for all data collectors. To capture the ground truth for step counts, each data collector also wore a pedometer and manually recorded his daily step count in the paper notebook. Each collector's pedometer was calibrated with his stride length and weight to obtain the most accurate step-count estimates possible.

We chose to use members of our research team to serve as data collectors because ground-truth diary logging is a tedious, error-prone process that required significant technical expertise to trouble-shoot problems with prototype technology. Given this overhead, the lack of application value to offer data collectors, and the high reliability of data logging that we required to test our algorithms, we felt that this was a reasonable choice. Our data collectors went to common places one would expect any person to visit such as grocery stores, malls, parks, churches, and libraries.

In all, the sensor logs contained 249 walking events (avg. 9.1 min) and 171 driving events (avg. 18.5 min). Each of these mobility events provides a sequence of data points to test our algorithm because every second is one data point to test our classifier (the rate that the phone scanned the radio environment). In total we gathered 12 GB of GSM network traces, amounting to 78 days of sensor logs. Our data spans urban and suburban environments and three different metropolitan areas as the data collectors traveled during the collection period.

3.2 Inferring Mobility Modes

Our goal was to infer one of three mobility states: *stationary*, *walking*, or *driving*. Periods of walking and driving were identified in the data collector's diaries. We had initially hoped to use the remaining times, which data collectors marked as being at a "place", to identify periods of being stationary. Unfortunately being at a "place" can still involve a fair degree of mobility. In a grocery store, shoppers are in motion much of the time. Even reasonably sedentary activities such as watching TV include short periods of walking (to visit the refrigerator for example). This ambiguity prevents us from having the needed ground truth for training and testing our algorithm. To extract the most reliable ground truth from our data, we used the GSM trace data collected between 2am and 5am to represent periods of being stationary. During these times we used our data collectors' logs to verify that they were at home and sleeping, thus their phones would not be moving. Although this means dropping much of our collected trace data, it provides the best possible ground truth for determining how well our classifier can differentiate properties of mobility.

Using the labeled periods of activity, we trained our classifier and evaluated it using a 5-fold cross validation[1] method over the entire data set. This produced a single model that worked well across all three data collectors and both GSM network providers. Figure 2 shows the precision, (true positive/(true positive + false positive), and recall, (true positive/(true positive + false negative), percentages aggregated for all of our data collectors. The percentages along the diagonal indicate the classifiers' performance for predicting and matching the ground truth events. Precision is the percentage of predicted events that are correct. A low precision number indicates many false positives. Recall is the percentage of ground truth events that were correctly identified. A low recall number indicates that many ground truth events were missed. Accuracy represents the percentage of predictions that are correct. Our overall accuracy, ((true positive + true negative)/ (total number of samples)), is 85%.

Our classification scheme performs very well for stationary periods correctly detecting most periods of no movement (recall 92.5%) and not raising many spurious stationary events (precision 95.4%). Driving also performs quite well detecting most drives (recall 81.7%) and not raising many false positives (precision 84.3%). Walking activities were also detected with high percentage (recall 80%), but exhibited the most false positives out of the three classes (precision 70.2%). Within a driving activity, there are often times when a car is moving at slow speeds such as in traffic or roads with lower speed limits. In our controlled experiment, we saw that the changes in signal strengths for slow driving speeds are similar to fast walking speeds. Thus, one would expect the classifier to predict walking movement even though a segment was

Precision

		Predicted Movement		
		Stationary	Walking	Driving
Ground Truth	Stationary	95.4%	12.6%	6.9%
	Walking	2.5%	70.2%	8.8%
	Driving	2.1%	17.2%	84.3%

Recall

		Predicted Movement		
		Stationary	Walking	Driving
Ground Truth	Stationary	92.5%	4.5%	3.0%
	Walking	7.7%	80.0%	12.2%
	Driving	4.5%	13.8%	81.7%

Fig. 2. Precision and recall confusion matrices for all GSM network traces aggregated over all data collectors. Overall accuracy is 85%.

1 In *k*-fold cross-validation, a data set is partitioned into *k*-folds, and *k* training and testing iterations are performed. On each iteration, *k-1* folds are used as a training set, and one fold is used as a testing set. The classification results from each iteration are averaged together to produce a final result.

marked as a driving activity. These types of misclassifications are reflected in the walking precision (17.2% driving) and driving recall (13.8% walking) numbers.

The results show that we are able to distinguish between different mobility states with high accuracy without having to instrument a person with any other additional sensors. The precision and recall numbers show that this type of scheme could be used in a person's daily life, to give an accurate diary of mobile activity. In Section 0 we will discuss several application domains where our techniques would be useful.

One question about our classification model is whether it is overfitted for our data set. As an external way to corroborate our classification model, we tested the model using the GSM traces gathered from our controlled experiment described in Section 0. These traces are independent of those used to build our model. The classifier achieved an overall accuracy of 90% on this controlled data set, with the only errors being that some portions of our slow walk were classified as stationary. Furthermore, boosting techniques have been shown to be robust to over fitting and generalizes well to unseen data [26].

3.3 Daily Step-Count Prediction for Data Collectors

To test the accuracy of a "virtual pedometer" capability, we asked our data collectors to wear an Omron Healthcare HJ-112 pedometer for a portion of the month during which they were collecting GSM data. We chose the Omron because it was rated as the overall best pedometer by Consumer Reports [9]. In all, we collected 50 days worth of daily step-count totals. In contrast to inferring mobility modes, for estimating step-count we want to be able to detect any walking activity throughout the day, even if it is for short periods of walking at a "place". The pedometer is always logging the steps a person takes, so our algorithm must also detect these periods of mobility. Thus, for step-count prediction we used all of the collected GSM trace data for each day.

We wanted our step-count predictor to work without any calibration for all users. This further allows us to promote ubiquitous mobility recognition with low setup costs. To predict a daily step count from our walking predictions, we used the following simple heuristic obtained by performing linear regression with a 5 fold cross validation on our data set:

$$\textit{daily step count} = 25 \cdot (\textit{minutes of walking})$$

For these 50 days of pedometer data, our heuristic predicted daily step counts ranging from 1500 to 12000 steps, with an average of 5000 steps. Comparing our estimates to the Omron step counts, we saw an average difference of 1400 steps per day (std. dev. 900 steps), with a minimum difference of 1 step and a maximum difference of 3500 steps. Our step count estimation worked uniformly well for all users: the correlation between measured and predicted step counts for the three data collectors were R=.71, .63, and .63. The error in our step count estimation is likely due more in part to errors in mobility estimation that to the user having different step rates.

To compare how well our step count predictions compared to other pedometers, we conducted a second experiment. We purchased four additional pedometers of varying brands, and collected seven more days of data for one data collector. For this experiment, he carried the GSM phone, while also wearing the Omron and the four other pedometers. Again, we used the Omron as ground truth in our evaluation. For these seven days, our GSM based predictions had an average difference of 1400 steps with a maximum difference of 2400. The average difference across the other pedometers varied between 500 and 900, with a maximum difference of 1500. These results show that while less accurate, our GSM-based step prediction approximates the results of off-the-shelf pedometers in predicting whether a person had a sedentary, moderately-active, or high-activity day.

4 Applications

Our mobility detection scheme provides a low-cost, ubiquitous method for high-level activity recognition. Since we use commodity GSM phones without any additional hardware, any owner of a GSM phone can use our mobility detection system. In this section, we describe two application domains where our mobility detection scheme would be useful.

4.1 Computer-Supported Coordinated Care (CSCC)

CSCC describes the network of people who help an elder *age in place*, *i.e.*, avoid the transition to a care facility, and seeks to improve the quality of her care while reducing the burden on the members of her care network, such as her family and friends [7]. The Digital Family Portrait [21] and CareNet Display [8] are two applications in the CSCC domain that aim to use sensor-driven activity inference to convey care and wellness information about an elder to members of her care network. The applications report information such as: Did the elder take her medication? Did she get out of bed? Did she have any visitors? Much research has focused on inferencing these types of in-home activities, but as the CareNet Display showed, an elder's care network is also concerned with activities that take place outside of the home, such as did the elder go to church on Sunday? Is she routinely late for her weekly doctor's appointment?

A recent report estimated that about 50% of Americans aged 65 to 74 are wireless customers and 30% of those aged 75 to 94 have mobile phones [2]. Given that so many elders already carry them, mobile phones present an interesting opportunity to provide detection of a range of activities that are meaningful to the elder's care network and can be detected today with a device that she already carries. With just a mobile phone, an elder would be able to relay information about her daily activity level, whether or not she was up and about today, or if she had a sedentary day around the house.

4.2 Social-Mobile Applications

Detecting mobility patterns is useful for applications that connect people with mobile devices together in their social environment--social mobile applications [28]. These applications-- if one includes voice calls and SMS --are key drivers of mobile phone usage today and are likely to continue as more and more of people's non-work lives revolve around mobile communications [16]. New applications are on the horizon that will help people communicate [30, 31] and coordinate [10, 29].

Mobility detection can provide context information to enhance these applications and provide a better experience for the user. For example, applications that prompt a user with information are competing for that person's attention and potentially interrupting an ongoing task. Our technique would be useful for example when driving, because the information might better serve the user if it is delayed. Mobility detection could be central to some applications such as one that computes estimated time of arrival for many people who want to rendezvous. In a scenario of this type, one user -- perhaps who is holding the movie tickets-- is very interested when the other 3 users will arrive. With mobility detection alone, the waiting user can discriminate that some others have parked already and are thus nearby and those who are still driving and thus distant; combining this with a location system provides an excellent tool for social coordination and obviates the need for many phone calls and SMS messages.

5 Related Work

The SHARP project aims to infer fine-grained activities by putting RFID tags on household objects and monitoring their usage with a wearable RFID reader [25]. Our approach complements the fine-grained activities SHARP can infer from instrumented objects, with high-level activities in the wider environment using low-resolution sensors.

GPS-based location sensing has been used for high-level activity recognition. Patterson et al. take a learning approach based on particle-filters to detect modes of transportation [23]. Similarly, Liao et al. extended Relational Markov Networks for learning models that, given a GPS location and the time, can differentiate among shopping, dining, visiting, at home, and at work [20]. GPS sensing today still often requires purchasing and carrying additional hardware. A recent study revealed that GPS positioning is available only about 5% of a typical person's day, as it needs a wide swath of clear sky to sense enough geostationary satellites [18]. In contrast, mobile phones provide ubiquitous coverage, and do not require any extra hardware from what people already carry. We have shown in this paper that similar recognition performance can be achieved observing changes in cell tower signal strengths, without the need for true location. This suggests that GPS should play an assistive role in everyday inference, rather than serving as the sole environmental sensor.

Two projects have looked at using radio signals for motion detection. LOCADIO used a Hidden Markov Model to infer motion of a device using 802.11 radio signals [17]. Anderson and Muller conducted a controlled, preliminary study with GSM mobile phones to detect motion of a device [4]. Similar to these two projects, our approach uses machine learning algorithms to infer motion. We have shown that motion detection using GSM is feasible for use outside the laboratory, and works well throughout people's daily lives.

A third approach to activity recognition is to use wearable sensors of a single modality [6] or multiple modalities [19]. Lester et al. use 7 different types of sensors, including light, audio, accelerometer, compass, temperature, humidity, and barometric pressure, to classify 10 activities such as sitting, standing, walking up stairs, and walking. The GSM radio can potentially be part of the sensor ensemble to improve recognition performance. Several commercial phones are now shipping with built-in accelerometers and compass, but, unfortunately, they do not expose the sensor readings to the application developers.

Finally, the Reality Mining project has used Bluetooth-capable GSM mobile phones to recognize social patterns in daily user activity, infer relationships, and model organizational rhythms [11]. It uses the single associated GSM cell tower, Bluetooth radio, application usage logs, and call logs to sense nearby Bluetooth phones and devices, time and duration of calls, caller ID, and so forth.

6 Conclusions and Future Work

We have demonstrated the feasibility of using an unmodified GSM phone, a coarse-grain but ubiquitous sensor with 1.5 billion subscribers worldwide [1], to recognize high-level properties of mobility that are valuable for application domains such as Computer-Supported Coordinated Care and social-mobile applications. To evaluate its effectiveness, we collected GSM traces and ground truth labels of walks and drives for a month from the everyday lives of three people, for a total of 78 days of GSM logs consisting of 249 walking events and 171 driving events. We have shown that we can recognize mobility modes among walking, driving, and stationary correctly 85% of the time, and estimate daily step counts that approximates commercial pedometers. Unlike other activity recognition systems that may require a person to wear a special device in a certain way, our approach lets users maintain their current mobile phone habits with no special requirements about where the phone is kept on their person. These results show that current mobile phones without extra sensors or devices can detect high-level activities, providing people with an estimate of their mobility patterns throughout the day.

Since our classification model was built mainly in one metropolitan area, we do not anticipate it working across different cell densities. However, building a model for our classifier with areas of different cell densities could enable our techniques to work in varying radio environments. Our future work involves exploring how our mobility detection technique and GSM-based step predictor would work in other parts of the country.

Acknowledgements

Thanks to Karen Tang who helped collect GSM network traces and provided feedback on several drafts of this paper. Susumu Harada, Dmitri Chmelev, and Fred Potter helped collect part of the data used for this research. Daniel Hsu explained many of the machine learning techniques used in this paper to us and gave several ideas on different methods to use. This research was supported in part by the UCSD FWGrid Project, NSF Research Infrastructure Grant EIA-0303622.

References

1. GSM Association Press Release, "Worldwide cellular connections exceeds 2 billion" http://www.gsmworld.com/news/press_2005/press05_21.shtml
2. Wireless Week, "Newer Phones, Older Users", http://www.wirelessweek.com/article/CA503601.html

3. Abowd, G. A. Bobick, I. Essa, E. Mynatt, and W. Rogers. The Aware Home: Developing Technologies for Successful Aging. In: Proceedings of AAAI Workshop and Automation as a Care Giver.
4. Anderson, I. and Muller, H. Context Awareness via GSM Signal Strength Fluctuation. In: Pervasive 2006, Late Breaking Results, pp. 27-31.
5. Bahl, P. and Padmanabhan, V. RADAR: An-In-Building RF-Based User Location and Tracking. In: Proceedings of IEEE Infocom 2000, pp. 775-784.
6. Bao, L. and Intille, S. S. Activity Recognition from User-Annotated Acceleration Data. In: Lecture Notes in Computer Science, Volume 3001, Jan 2004, Pages 1-17
7. Consolvo, S., Roessler, P., Shelton, B.E., LaMarca, A., Schilit, B., Bly, S. Technology for Care Networks of Elders. IEEE Pervasive Computing Mobile & Ubiquitous Systems: Successful Aging, Vol. 3, No. 2 (Apr-June 2004) pp. 22-29.
8. Consolvo, S., Roessler, P., and Shelton, B.E. The CareNet Display: Lessons Learned from an In Home Evaluation of an Ambient Display. In: Proceedings of UbiComp 2004, pp. 1-17.
9. Consumer Reports, "Pedometers: Walking by the numbers," Consumer Reports, 69(10), (2004), pp.30-2.
10. Dodgeball. http://www.dodgeball.com/
11. Eagle, N. and Pentland, A. Reality Mining: Sensing Complex Social Systems. Personal and Ubiquitous Computing (Jan 2006).
12. Freund, Y. and Schapire, R.E. Experiments with a new boosting algorithm. In: Proceedings of International Conference on Machine Learning, pp. 148-156, 1996.
13. Friedman, J.H., Hastie, T., Tibshirani, R. Additive logistic regression: A statistical view of boosting. Annals of Statistics (2000) 337--374
14. Hariharan, R., Krumm, J., Horvitz, E. Web-Enhanced GPS. In: Proceedings of the International Workshop on Location and Context Awareness (LoCA 2005), May 2005.
15. Intille, S. S., Larson, K., Beaudin, J. S., Nawyn, J., Tapia, E. M., and Kaushik, P. A living laboratory for the design and evaluation of ubiquitous computing technologies. In: CHI 2005 Extended Abstracts, pp. 1941-1944.
16. Ito, Mizuko, et. Al., eds. Personal, Portable, Pedestrian. Mobile Phones in Japanese Life. Cambridge, The MIT Press, 2005.
17. Krumm, J. and Horvitz, E. LOCADIO: Inferring Motion and Location from Wi-Fi Signal Strengths. In: Mobiquitous 2004, pp. 4-13.
18. LaMarca, A., Chawathe, Y., Consolvo, S., Hightower, J., Smith, I., Scott, J., Sohn, T., Howard, J., Hughes, J., Potter, F., Tabert, J., Powledge, P., Borriello, G., Schilit, B. Place Lab: Device Positioning Using Radio Beacons in the Wild. In: Pervasive 2005.
19. Lester, J., Choudhury, T., Kern, N., Borriello, G., and Hannaford, B. A Hybrid Discriminative-Generative Approach for Modeling Human Activities. In: Proceedings of International Joint Conference on Artificial Intelligence (IJCAI 2005).
20. Liao, L., Fox, D., Kautz, H. Location-Based Activity Recognition using Relational Markov Networks. In: Proceedings of the International Conference on Artificial Intelligence (IJCAI 2005).
21. Mynatt, E. D., Rowan, J., Craighill, S., and Jacobs. Digital family portraits: supporting peace of mind for extended family members. In: Proceedings of the CHI 2001, pp. 333-340.
22. Otsason, V., Varshavsky, A., LaMarca, A., and de Lara, E. Accurate GSM Indoor Localization. In: Ubicomp 2005.
23. Patterson, D.J., Liao, L., Fox, D., Kautz. Inferring High-Level Behavior from Low-Level Sensors. In: Ubicomp 2003.

24. Pavlovic, V., Garg, A., and Rehg, J.M. Multimodal speaker detection using error feedback dynamic bayesian networks. In: IEEE Conference on Computer Vision and Pattern Recognition 2000.
25. Philipose, M., Fishkin, K., Perkowitz, M., Patterson, D.J., Fox, D., Kautz, H., Hähnel, D., Inferring Activities from Interactions with Objects. In: IEEE Pervasive Computing, Oct 2004.
26. Schapire, R.E. The boosting approach to machine learning: An overview. In: D. D. Denison, M. H. Hansen, C. Holmes, B. Mallick, B. Yu, editors, Nonlinear Estimation and Classification. Springer, 2003
27. Schwenk, H. and Bengio, Y. Training methods for adaptive boosting of neural networks for character recognition. In: Proceedings of NIPS'98.
28. Smith, I. Social-Mobile Applications. IEEE Computer 38(4): 84-85 (2005)
29. Smith, I., Consolvo, S., Hightower, J., Hughes, J., Iachello, G., LaMarca, A., Scott, J., Sohn, T., Abowd, G. Social Disclosure of Place: From Location Technology to Communication Practice. In: Pervasive 2005.
30. Socialight. http://socialight.com
31. Textamerica. http://www.textamerica.com/
32. Tudor-Locke C, Bassett DR Jr., "How many steps/day are enough? Preliminary pedometer indices for public health," Sports Med., 34(1), (2004), pp.1-8.
33. Viola, P.A. and Jones, M.J. Robust real-time face detection. In: International Journal of Computer Vision, vol. 57, no. 2, pp. 137--154, 2004.
34. Ward, A. Jones, A. Hopper, A. A new location technique for the active office. In: Personal Communications. Oct 1997.
35. Webb, G.I. MultiBoosting: A Technique for Combining Boosting and Wagging. Machine Learning, 40(2): 159-196, 2000.
36. Weisstein, E.W. "Spearman Rank Correlation Coefficient." http://mathworld.wolfram. com/SpearmanRankCorrelationCoefficient.html
37. Witten, I.H. and Frank, E. "Data Mining: Practical machine learning tools and techniques", 2nd Edition, Morgan Kaufmann, San Francisco, 2005.