Activity Recognition using Cell Phone Accelerometers

Jennifer R. Kwapisz, Gary M. Weiss, Samuel A. Moore

Department of Computer and Information Science
Fordham University
441 East Fordham Road
Bronx, NY 10458
{kwapisz, gweiss, asammoore}@cis.fordham.edu

ABSTRACT

Mobile devices are becoming increasingly sophisticated and the latest generation of smart cell phones now incorporates many diverse and powerful sensors. These sensors include GPS sensors, vision sensors (i.e., cameras), audio sensors (i.e., microphones), light sensors, temperature sensors, direction sensors (i.e., magnetic compasses), and acceleration sensors (i.e., accelerometers). The availability of these sensors in mass-marketed communication devices creates exciting new opportunities for data mining and data mining applications. In this paper we describe and evaluate a system that uses phone-based accelerometers to perform *activity recognition*, a task which involves identifying the physical activity a user is performing. To implement our system we collected labeled accelerometer data from twenty-nine users as they performed daily activities such as walking, jogging, climbing stairs, sitting, and standing, and then aggregated this time series data into examples that summarize the user activity over 10-second intervals. We then used the resulting training data to induce a predictive model for activity recognition. This work is significant because the activity recognition model permits us to gain useful knowledge about the habits of millions of users passively—just by having them carry cell phones in their pockets. Our work has a wide range of applications, including automatic customization of the mobile device's behavior based upon a user's activity (e.g., sending calls directly to voicemail if a user is jogging) and generating a daily/weekly activity profile to determine if a user (perhaps an obese child) is performing a healthy amount of exercise.

Categories and Subject Descriptors

I.2.6 [**Artificial Intelligence**]: Learning-*induction*

General Terms

Algorithms, Design, Experimentation, Human Factors

Keywords

Sensor mining, activity recognition, induction, cell phone, accelerometer, sensors

1. INTRODUCTION

Mobile devices, such as cellular phones and music players, have recently begun to incorporate diverse and powerful sensors. These sensors include GPS sensors, audio sensors (i.e., microphones), image sensors (i.e., cameras), light sensors, temperature sensors, direction sensors (i.e., compasses) and acceleration sensors (i.e., accelerometers). Because of the small size of these "smart" mobile devices, their substantial computing power, their ability to send and receive data, and their nearly ubiquitous use in our society, these devices open up exciting new areas for data mining research and data mining applications. The goal of our WISDM (Wireless Sensor Data Mining) project [19] is to explore the research issues related to mining sensor data from these powerful mobile devices and to build useful applications. In this paper we explore the use of one of these sensors, the accelerometer, in order to identify the activity that a user is performing—a task we refer to as activity recognition.

We have chosen Android-based cell phones as the platform for our WISDM project because the Android operating system is free, open-source, easy to program, and expected to become a dominant entry in the cell phone marketplace (this is clearly happening). Our project currently employs several types of Android phones, including the Nexus One, HTC Hero, and Motorola Backflip. These phones utilize different cellular carriers, although this is irrelevant for our purposes since all of the phones can send data over the Internet to our server using a standard interface. However, much of the data in this work was collected directly from files stored on the phones via a USB connection, but we expect this mode of data collection to become much less common in future work.

All of these Android phones, as well as virtually all new smart phones and smart music players, including the iPhone and iPod Touch [2], contain tri-axial accelerometers that measure acceleration in all three spatial dimensions. These accelerometers are also capable of detecting the orientation of the device (helped by the fact that they can detect the direction of Earth's gravity), which can provide useful information for activity recognition. Accelerometers were initially included in these devices to support advanced game play and to enable automatic screen rotation but they clearly have many other applications. In fact, there are many useful applications that can be built if accelerometers can be used to recognize a user's activity. For example, we can automatically monitor a user's activity level and generate daily, weekly, and monthly activity reports, which could be automatically emailed to the user. These reports would indicate an overall activity level,

which could be used to gauge if the user is getting an adequate amount of exercise and estimate the number of daily calories expended. These reports could be used to encourage healthy practices and might alert some users to how sedentary they or their children actually are. The activity information can also be used to automatically customize the behavior of the mobile phone. For example, music could automatically be selected to match the activity (e.g., "upbeat" music when the user is running) or send calls directly to voicemail when the user is exercising. There are undoubtedly numerous other instances where it would be helpful to modify the behavior of the phone based on the user activity and we expect that many such applications will become available over the next decade.

In order to address the activity recognition task using supervised learning, we first collected accelerometer data from twenty-nine users as they performed activities such as walking, jogging, ascending stairs, descending stairs, sitting, and standing. We then aggregated this raw time series accelerometer data into examples, as described in Section 2.2, where each example is labeled with the activity that occurred while that data was being collected. We then built predictive models for activity recognition using three classification algorithms.

The topic of accelerometer-based activity recognition is not new. Bao & Intille [3] developed an activity recognition system to identify twenty activities using bi-axial accelerometers placed in five locations on the user's body. Additional studies have similarly focused on how one can use a variety of accelerometer-based devices to identify a range of user activities [4-7, 9-16, 21]. Other work has focused on the applications that can be built based on accelerometer-based activity recognition. This work includes identifying a user's activity level and predicting their energy consumption [8], detecting a fall and the movements of user after the fall [12], and monitoring user activity levels in order to promote health and fitness [1]. Our work differs from most prior work in that we use a commercial mass-marketed device rather than a research-only device, we use a single device conveniently kept in the user's pocket rather than multiple devices distributed across the body, and we require no additional actions by the user. Also, we have generated and tested our models using more users (twenty-nine) than most previous studies and expect this number to grow substantially since we are continuing to collect data. The few studies that have involved commercial devices such as smart phones have focused either on a very small set of users [21] or have trained models for particular users [4] rather than creating a universal model that can be applied to any user.

Our work makes several contributions. One contribution is the data that we have collected and continue to collect, which we plan to make public in the future. This data can serve as a resource to other researchers, since we were unable to find such publically available data ourselves. We also demonstrate how raw time series accelerometer data can be transformed into examples that can be used by conventional classification algorithms. We demonstrate that it is possible to perform activity recognition with commonly available (nearly ubiquitous) equipment and yet achieve highly accurate results. Finally, we believe that our work will help bring attention to the opportunities available for mining wireless sensor data and will stimulate additional work in this area.

The remainder of this paper is structured as follows. Section 2 describes the process for addressing the activity recognition task, including data collection, data preprocessing, and data transformation. Section 3 describes our experiments and results. Related work is described in Section 4 and Section 5 summarizes our conclusions and discusses areas for future research.

2. THE ACTIVITY RECOGNITION TASK
In this section we describe the activity recognition task and the process for performing this task. In Section 2.1 we describe our protocol for collecting the raw accelerometer data, in Section 2.2 we describe how we preprocess and transform the raw data into examples, and in Section 2.3 we describe the activities that will be predicted/identified.

2.1 Data Collection
In order to collect data for our supervised learning task, it was necessary to have a large number of users carry an Android-based smart phone while performing certain everyday activities. Before collecting this data, we obtained approval from the Fordham University IRB (Institutional Review Board) since the study involved "experimenting" on human subjects and there was some risk of harm (e.g., the subject could trip while jogging or climbing stairs). We then enlisted the help of twenty-nine volunteer subjects to carry a smart phone while performing a specific set of activities. These subjects carried the Android phone in their front pants leg pocket and were asked to walk, jog, ascend stairs, descend stairs, sit, and stand for specific periods of time.

The data collection was controlled by an application we created that executed on the phone. This application, through a simple graphical user interface, permitted us to record the user's name, start and stop the data collection, and label the activity being performed. The application permitted us to control what sensor data (e.g., GPS, accelerometer) was collected and how frequently it was collected. In all cases we collected the accelerometer data every 50ms, so we had 20 samples per second. The data collection was supervised by one of the WISDM team members to ensure the quality of the data.

2.2 Feature Generation & Data Transformation
Standard classification algorithms cannot be directly applied to raw time-series accelerometer data. Instead, we first must transform the raw time series data into examples [18]. To accomplish this we divided the data into 10-second segments and then generated features that were based on the 200 readings contained within each 10-second segment. We refer to the duration of each segment as the example duration (ED). We chose a 10-second ED because we felt that it provided sufficient time to capture several repetitions of the (repetitive) motions involved in some of the six activities. Although we have not performed experiments to determine the optimal example duration value, we did compare the results for a 10-second and 20-second ED and the 10-second ED yielded slightly better results (as well as twice as many training examples).

Next we generated informative features based on the 200 raw accelerometer readings, where each reading contained an x, y, and z value corresponding to the three axes/dimensions (see Figure 1). We generated a total of forty-three summary features, although these are all variants of just six basic features. The features are

described below, with the number of features generated for each feature-type noted in brackets:

- Average[3]: Average acceleration (for each axis)

- Standard Deviation[3]: Standard deviation (for each axis)

- Average Absolute Difference[3]: Average absolute difference between the value of each of the 200 readings within the ED and the mean value over those 200 values (for each axis)

- Average Resultant Acceleration[1]: Average of the square roots of the sum of the values of each axis squared $\sqrt{(x_i^2 + y_i^2 + z_i^2)}$ over the ED

- Time Between Peaks[3]: Time in milliseconds between peaks in the sinusoidal waves associated with most activities (for each axis)

- Binned Distribution[30]: We determine the range of values for each axis (maximum – minimum), divide this range into 10 equal sized bins, and then record what fraction of the 200 values fell within each of the bins.

The "time between peaks" feature requires further explanation. The repetitive activities, like walking, tend to generate repeating waves for each axis and this feature tries to measure the time between successive peaks. To estimate this value, for each example we first identify all of the peaks in the wave using a heuristic method and then identify the highest peak for each axis. We then set a threshold based on a percentage of this value and find the other peaks that met or exceed this threshold; if no peaks meet this criterion then the threshold is lowered until we find at least three peaks. We then measure the time between successive peaks and calculate the average. For samples where at least three peaks could not be found, the time between peaks is marked as unknown. This method was able to accurately find the time between peaks for the activities that had a clear repetitive pattern, like walking and jogging. Certainly more sophisticated schemes will be tried in the future.

The number of examples generated per user for each activity varies. These differences are due to the time limitations that some users may have or physical limitations that impact the time they spend on each activity. Our data set is summarized in Section 3.1.

2.3 The Activities

In this study we consider six activities: walking, jogging, ascending stairs, descending stairs, sitting, and standing. We selected these activities because they are performed regularly by many people in their daily routines. The activities also involve motions that often occur for substantial time periods, thus making them easier to recognize. Furthermore, most of these activities involve repetitive motions and we believe this should also make the activities easier to recognize. When we record data for each of these activities, we record acceleration in three axes. For our purposes, the z-axis captures the forward movement of the leg and the y-axis captures the upward and downward motion. The x-axis captures horizontal movement of the user's leg. Figure 1 demonstrates these axes relative to a user.
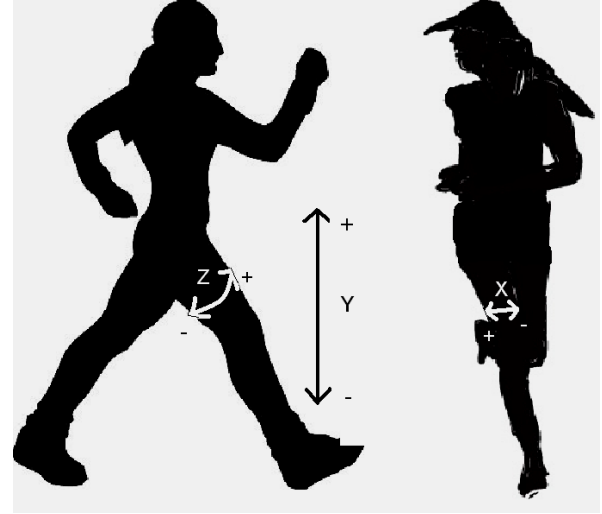

**Figure 1: Axes of Motion Relative to User**

Figure 2 plots the accelerometer data for a typical user, for all three axes and for each of the six activities. It is clear that sitting and standing (Figure 2e,f) do not exhibit periodic behavior but do have distinctive patterns, based on the relative magnitudes of the x, y, and z, values, while the four other activities (Figure 2a-d), which involve repetitive motions, do exhibit periodic behavior. Note that for most activities the y values have the largest accelerations. This is a consequence of Earth's gravitational pull, which causes the accelerometer to measure a value of 9.8 m/s^2 in the direction of the Earth's center. For all activities except sitting this direction corresponds to the y axis (see Figure 1).

The periodic patterns for walking, jogging, ascending stairs, and descending stairs (Figure 2a-d) can be described in terms of the time between peaks and by the relative magnitudes of the acceleration values. The plot for walking, shown in Figure 2a, demonstrates a series of high peaks for the y-axis, spaced out at approximately ½ second intervals. The peaks for the z-axis acceleration data echo these peaks but with a lower magnitude. The distance between the peaks of the z-axis and y-axis data represent the time of one stride. The x-axis values (side to side) have an even lower magnitude but nonetheless mimic the peaks associated with the other axes. For jogging, similar trends are seen for the z-axis and y-axis data, but the time between peaks is less (~¼ second), as one would expect. As one might expect, the range of y-axis acceleration values for jogging is greater than for walking, although the shift is more noticeable in the negative direction.

For descending stairs, one observes a series of small peaks for y-axis acceleration that take place every ~½ second. Each small peak represents movement down a single stair. The z-axis values show a similar trend with negative acceleration, reflecting the regular movement down each stair. The x-axis data shows a series of semi-regular small peaks, with acceleration vacillating again between positive and negative values. For ascending stairs, there are a series of regular peaks for the z-axis data and y-axis data as well; these are spaced approximately ~¾ seconds apart, reflecting the longer time it takes to climb up stairs.
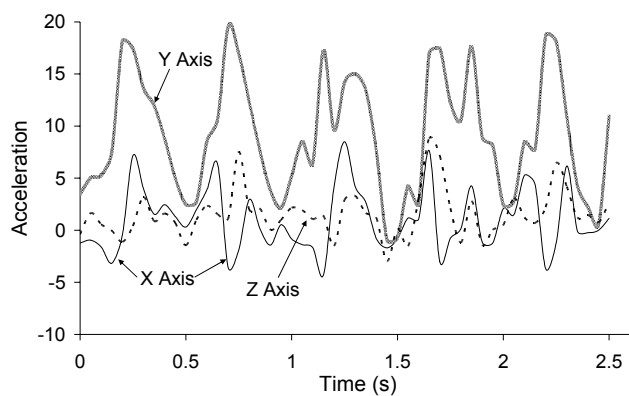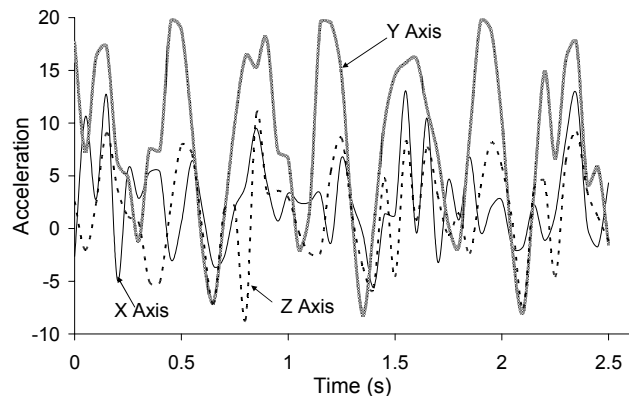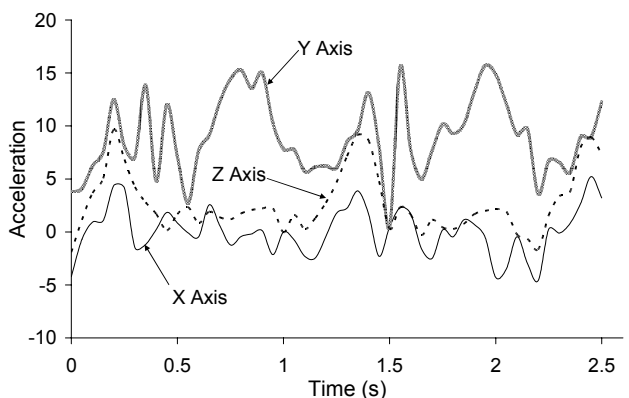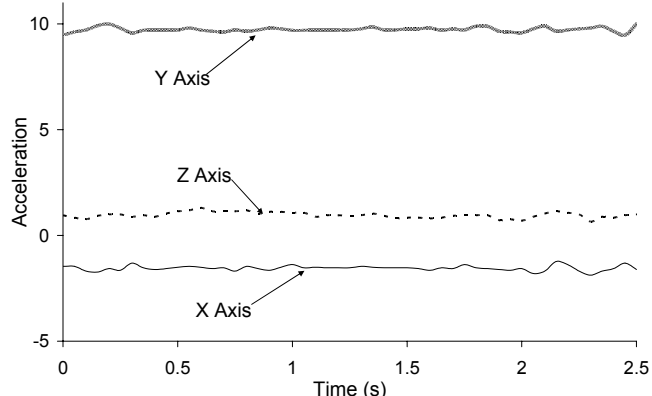
Figure 2: Acceleration Plots for the Six Activities (a-f)

As one would expect, sitting and standing do not exhibit any regular periodic behavior and all of the acceleration values are relatively constant. As mentioned earlier, the primary differences between 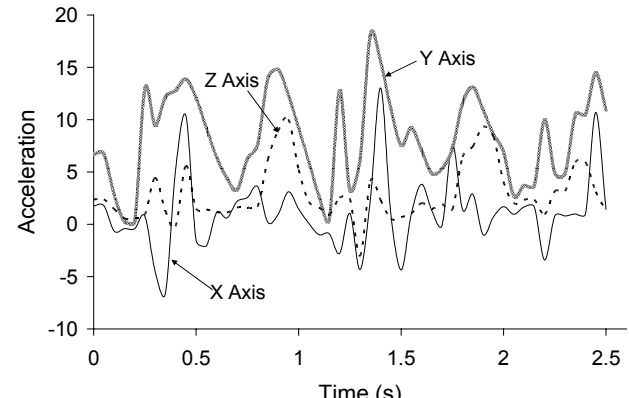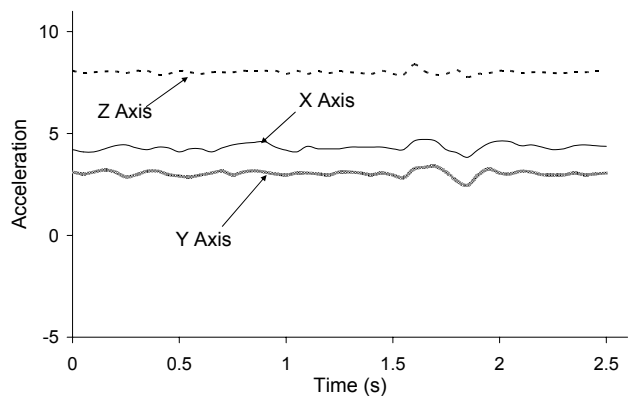these activities is the relative magnitudes of values for each axis, due to the different orientations of the device with respect to the Earth when the user is sitting and standing Thus it appears easy to differentiate between sitting and standing, even though neither involves much movement. Note that because the accelerometers are themselves able to determine orientation with respect to the Earth's gravitational field, it would be relatively straightforward to compensate/correct for any changes in the cell phone's orientation due to the phone shifting position in a user's pocket. We plan to implement this correction in future work.

3. EXPERIMENTS

In this section we describe our experiments and then present and discuss our results for the activity recognition task.

3.1 Description of Experiments

Our experiments first require us to collect the labeled raw accelerometer data and then transform that data into examples. This process was described in Section 2. The resulting examples contain 43 features and cover twenty-nine users. This forms the data set, described in Table 1, which is subsequently used for training and testing. The last row in Table 1 shows the percentage of the total examples associated with each activity.

Table 1: Number of Examples per User and Activity

ID	Walk	Jog	Up	Down	Sit	Stand	Total
1	74	15	13	25	17	7	151
2	48	15	30	20	0	0	113
3	62	58	25	23	13	9	190
4	65	57	25	22	6	8	183
5	65	54	25	25	77	27	273
6	62	54	16	19	11	8	170
7	61	55	13	11	9	4	153
8	57	54	12	13	0	0	136
9	31	59	27	23	13	10	163
10	62	52	20	12	16	9	171
11	64	55	13	12	8	9	161
12	36	63	0	0	8	6	113
13	60	62	24	15	0	0	161
14	62	0	7	8	15	10	102
15	61	32	18	18	9	8	146
16	65	61	24	20	0	8	178
17	70	0	15	15	7	7	114
18	66	59	20	20	0	0	165
19	69	66	41	15	0	0	191
20	31	62	16	15	4	3	131
21	54	62	15	16	12	9	168
22	33	61	25	10	0	0	129
23	30	5	8	10	7	0	60
24	62	0	23	21	8	15	129
25	67	64	21	16	8	7	183
26	85	52	0	0	14	17	168
27	84	70	24	21	11	13	223
28	32	19	26	22	8	15	122
29	65	55	19	18	8	14	179
Sum	1683	1321	545	465	289	223	4526
%	37.2	29.2	12.0	10.2	6.4	5.0	100

Note that certain activities contain fewer examples than others, mainly because the users were not asked to perform strenuous activities (e.g., jogging, climbing stairs) for very long and because we thought that the patterns in other activities (e.g., standing) would become apparent quickly so that there would be no need to waste the users time literally "standing around." Furthermore, certain activities, like standing and sitting, were only added after the study began, so we have no data for these activities for some users.

Once the data set was prepared, we used three classification techniques from the WEKA data mining suite [20] to induce models for predicting the user activities: decision trees (J48), logistic regression and multilayer neural networks. In each case we used the default settings. We used ten-fold cross validation for all experiments and all results are based on these ten runs.

3.2 Results

The summary results for our activity recognition experiments are presented in Table 2. This table specifies the predictive accuracy associated with each of the activities, for each of the three learning algorithms and for a simple "straw man" strategy. The straw man strategy always predicts the specified activity (i.e., walking for the first row in Table 2 and jogging for the second row of Table 2) or, when assessing the overall performance of the classifier (i.e., the last row of Table 2), always predicts the most frequently occurring activity, which happens to be walking. The baseline straw man strategy allows us to consider the degree of class imbalance when evaluating the performance of the activity recognition system.

Table 2: Accuracies of Activity Recognition

	% of Records Correctly Predicted			
	J48	**Logistic Regression**	**Multilayer Perceptron**	**Straw Man**
Walking	89.9	**93.6**	91.7	37.2
Jogging	96.5	98.0	**98.3**	29.2
Upstairs	59.3	27.5	**61.5**	12.2
Downstairs	**55.5**	12.3	44.3	10.0
Sitting	**95.7**	92.2	95.0	6.4
Standing	**93.3**	87.0	91.9	5.0
Overall	85.1	78.1	**91.7**	37.2

Table 2 demonstrates that in most cases we can achieve high levels of accuracy. For the two most common activities, walking and jogging, we generally achieve accuracies above 90%. Jogging appears easier to identify than walking, which seems to make sense, since jogging involves more extreme changes in acceleration. It appears much more difficult to identify the two stair climbing activities, but as we shall see shortly, that is because those two similar activities are often confused with one another. Note that although there are very few examples of sitting and standing, we can still identify these activities quite well, because, as noted earlier, the two activities cause the device to change orientation and this is easily detected from the accelerometer data. Our results indicate that none of the three learning algorithms consistently performs best, but the multilayer perceptron does perform best overall. More detailed results are presented in Tables

3-5, which show the confusion matrices associated with each of the three learning algorithms.

Table 3: Confusion Matrix for J48

		Predicted Class					
		Walk	Jog	Up	Down	Sit	Stand
Actual Class	Walk	**1513**	14	72	82	2	0
	Jog	16	**1275**	16	12	1	1
	Up	88	23	**323**	107	2	2
	Down	99	13	92	**258**	1	2
	Sit	4	0	2	3	**270**	3
	Stand	4	1	2	7	1	**208**

Table 4: Confusion Matrix for Logistic Regression

		Predicted Class					
		Walk	Jog	Up	Down	Sit	Stand
Actual Class	Walk	**1575**	14	53	36	2	3
	Jog	15	**1294**	6	6	0	0
	Up	277	36	**150**	77	1	4
	Down	259	6	136	**57**	3	4
	Sit	1	0	4	11	**260**	6
	Stand	3	1	7	3	15	**194**

Table 5: Confusion Matrix for Multilayer Perceptron

		Predicted Class					
		Walk	Jog	Up	Down	Sit	Stand
Actual Class	Walk	**1543**	5	73	60	1	1
	Jog	3	**1299**	16	3	0	0
	Up	84	24	**335**	98	2	2
	Down	108	10	136	**206**	2	3
	Sit	0	2	4	1	**268**	7
	Stand	1	0	5	4	8	**205**

The most important activities to analyze are the climbing-up and climbing-down stair activities, since these were the only activities that that were difficult to recognize. The confusion matrices indicate that many of the prediction errors are due to confusion between these two activities. If we focus on the results for the J48 decision tree model in Table 3, we see that when we are climbing up stairs the most common incorrect classification occurs when we predict "downstairs," which occurs 107 times and accounts for a decrease in accuracy of 19.6% (107 errors out of 545). When the actual activity is climbing downstairs, walking slightly outpaces "upstairs" in terms of the total number of errors (99 vs. 92), but this is only because walking occurs more than three times as often as climbing upstairs in our dataset. If we look at Figures 2a, 2c, and 2d, we see that the patterns in acceleration data between "walking", "ascending stairs" and "descending stairs" are somewhat similar. To limit the confusion between the ascending and descending stair activities, we ran another set of experiments where we combine ascending stairs and descending stairs into one activity. The resulting confusion matrix for the J48 algorithm is shown in Table 6 (in the interest of space we do not show them for the other two algorithms). We see that the results are substantially improved, although stair climbing is still the hardest activity to recognize.

Table 6: Confusion Matrix for J48 Model (Stairs Combined)

		Predicted Class					Accur. (%)
		Walk	Jog	Stairs	Sit	Stand	
Actual Class	Walk	**1524**	7	148	2	2	90.6
	Jog	10	**1280**	31	0	0	96.9
	Stairs	185	33	**784**	4	4	<u>77.6</u>
	Sit	4	0	2	**272**	4	96.5
	Stand	3	1	10	0	**209**	93.7

4. RELATED WORK

Activity recognition has recently gained attention as a research topic because of the increasing availability of accelerometers in consumer products, like cell phones, and because of the many potential applications. Some of the earliest work in accelerometer-based activity recognition focused on the use of multiple accelerometers placed on several parts of the user's body. In one of the earliest studies of this topic, Bao & Intille [3] used five biaxial accelerometers worn on the user's right hip, dominant wrist, non-dominant upper arm, dominant ankle, and non-dominant thigh in order to collect data from 20 users. Using decision tables, instance-based learning, C4.5 and Naïve Bayes classifiers, they created models to recognize twenty daily activities. Their results indicated that the accelerometer placed on the thigh was most powerful for distinguishing between activities. This finding supports our decision to have our test subjects carry the phone in the most convenient location—their pants pocket.

Other researchers have, like Bao & Intille, used multiple accelerometers for activity recognition. Krishnan et. al. [9] collected data from three users using two accelerometers to recognize five activities—walking, sitting, standing, running, and lying down. This paper claimed that data from a thigh accelerometer was insufficient for classifying activities such as sitting, lying down, walking, and running, and thus multiple accelerometers were necessary (a claim our research contradicts). In another paper, Krishnan et. al. [10] examined seven lower body activities using data collected from ten subjects wearing three accelerometers. This method was tested in supervised and semi-naturalistic settings. Tapia et. al. [16] collected data from five accelerometers placed on various body locations for twenty-one users and used this data to implement a real-time system to recognize thirty gymnasium activities. A slight increase in performance was made by incorporating data from a heart monitor in addition to the accelerometer data. Mannini and Sabitini [23] used five tri-axial accelerometers attached to the hip, wrist, arm, ankle, and thigh in order to recognize twenty activities from thirteen users. Various learning methods were used to recognize three "postures" (lying, sitting, and standing) and five "movements" (walking, stair climbing, running, and cycling). Foerster and Fahrenberg [28] used data from five accelerometers in one set of experiments and from two of those accelerometers in another for activity recognition. Thirty-one male subjects participated in the study and a hierarchical classification model was built in order to distinguish between postures such as sitting and lying at specific angles, and motions such as walking and climbing stairs at different speeds.

Researchers have used a combination of accelerometers and other sensors to achieve activity recognition. Parkka et. al. [27] created

a system using twenty different types of sensors (including an accelerometer worn on the chest and one worn on the wrist) in order to recognize activities such as lying, standing, walking, running, football, swinging, croquet, playing ball, and using the toilet in specific locations. Lee and Mase [25] created a system to recognize a user's location and activities, including sitting, standing, walking on level ground, walking upstairs, and walking downstairs using a sensor module that consisted of a biaxial accelerometer and an angular velocity sensor worn in the pocket combined with a digital compass worn at the user's waist. Subramanya et. al. [26] addressed similar activities by building a model using data from a tri-axial accelerometer, two microphones, phototransitors, temperature and barometric pressure sensors, and GPS to distinguish between a stationary state, walking, jogging, driving a vehicle, and climbing up and down stairs.

While these systems using multiple accelerometers or a combination of accelerometers and other sensors were capable of identifying a wide range of activities, they are not very practical because they involve the user wearing multiple sensors distributed across their body. This could work for some short term, small scale, highly specialized applications (e.g., in a hospital setting) but would certainly not work for the applications that we envision.

Some studies have also focused on combining multiple types of sensors in addition to accelerometers for activity recognition. Maurer et al. [13] used "eWatch" devices placed on the belt, shirt pocket, trouser pocket, backpack, and neck to recognize the same six activities that we consider in our study. Each "eWatch" consisted of a biaxial accelerometer and a light sensor. Decision trees, k-Nearest Neighbor, Naïve Bayes, and Bayes Net classifiers with five-fold cross validation were used for learning. Choudhury et. al [6] used a multimodal sensor device consisting of seven different types of sensors (tri-axial accelerometer, microphone, visible light phototransitor, barometer, visible+IR light sensor, humidity/temperature reader, and compass) to recognize activities such as walking, sitting, standing, ascending stairs, descending stairs, elevator moving up and down, and brushing one's teeth. Cho et. al. [5] used a single tri-axial accelerometer, along with an embedded image sensor worn at the user's waist, to identify nine activities. Although these multi-sensor approaches do indicate the great potential of mobile sensor data as more types of sensors are being incorporated into devices, our approach shows that only one type of sensor—an accelerometer—is needed to recognize most daily activities. Thus our method offers a straightforward and easily-implementable approach to accomplish this task.

Other studies, like our own, have focused on the use of a single accelerometer for activity recognition. Long, Yin, and Aarts [22] collected accelerometer data from twenty-four users using a tri-axial accelerometer worn without regard for orientation at the user's waist. Data was collected naturalistically, and decision trees as well as a Bayes classifier combined with a Parzen window estimator were used to recognize walking, jogging, running, cycling, and sports. Lee et. al. [24] used a single accelerometer attached to the left waists of five users. Standing, sitting, walking, lying, and running were all recognized with high accuracies using fuzzy c-means classification. However unlike these studies, which use devices specifically made for research purposes, our method utilizes commercial devices that are widely-available without any additional specialized equipment. This approach enables make practical real-world applications for our models.

Several researchers have considered the use of widely-available mobile devices such as cell phones to address the activity recognition problem. However the earlier approaches did not take advantage of the sensors incorporated into the mobile devices themselves. For example, Gyorbiro et. al. [7] used "MotionBands" attached to the dominant wrist, hip, and ankle of each subject to distinguished between six different motion patterns. Each MotionBand contained a tri-axial accelerometer, magnetometer, and gyroscope. As the MotionBand collected data, the data was then transmitted to a smart phone carried by the user to be stored. Ravi et. al. [15] collected data from two users wearing a single accelerometer-based device and then transmitted this data to the HP iPAQ mobile device carried by the user. Using this data for activity recognition, researchers compared the performance of eighteen different classifiers. Lester et. al. [11] used accelerometer data, along with audio and barometric sensor data, to recognize eight daily activities from a small set of users. While these studies could have used a cell phone to generate the accelerometer data, they did not do this. Instead, the data was generated using distinct accelerometer-based devices worn by the user and then sent to a cell phone for storage.

A few studies, like ours, did use an actual commercial mobile device to collect data for activity recognition. Such systems offer an advantage over other accelerometer-based systems because they are unobtrusive and do not require any additional equipment for data collection and accurate recognition. Miluzzo et. al. [14] explored the use of various sensors (such as a microphone, accelerometer, GPS, and camera) available on commercial smart phones for activity recognition and mobile social networking applications. In order to address the activity recognition task, they collected accelerometer data from ten users to build an activity recognition model for walking, running, sitting, and standing using J48. This model had particular difficulty distinguishing between the sitting and standing activities, a task that our models easily achieve. Yang [21] developed an activity recognition system using the Nokia N95 phone to distinguish between sitting, standing, walking, running, driving, and bicycling. This work also explored the use of an activity recognition model to construct physical activity diaries for the users. Although the study achieved relatively high accuracies of prediction, stair climbing was not considered and the system was trained and tested using data from only four users. Brezmes et. al. [4] also used the Nokia N95 phone to develop a real-time system for recognizing six user activities. In their system, an activity recognition model is trained for each user, meaning that there is no universal model that can be applied to new users, for whom no training data exists. Our models do not have this limitation.

5. CONCLUSIONS AND FUTURE WORK

In this paper we described how a smart phone can be used to perform activity recognition, simply by keeping it in ones pocket. We further showed that activity recognition can be highly accurate, with most activities being recognized correctly over 90% of the time. In addition, these activities can be recognized quickly, since each example is generated from only 10 seconds worth of data. We have several interesting applications in mind for activity recognition and plan to implement some of these applications in the near future.

Our work would not have been possible without establishing our WISDM Android-based data collection platform, and we view

this software and hardware architecture, where data is transmitted by the phone to our Internet-based server, as a key resource produced as a consequence of this work. By having this in place we will be able to mine other mobile sensor data much more quickly. This platform, as well as the data that we collected, will ultimately be made public.

We plan to improve our activity recognition in several ways. The straightforward improvements involve: 1) learning to recognize additional activities, such as bicycling and car-riding, 2) obtaining training data from more users with the expectation that this will improve our results, 3) generating additional and more sophisticated features when aggregating the raw time-series data, and 4) evaluating the impact of carrying the cell phone in different locations, such as on a belt loop. In addition, in the near future we plan to significantly enhance our WISDM platform so that we can generate results in real-time, whereas currently our results are generated off-line and are not reported back to the mobile phone and the user. We plan to provide real-time results in two ways. The first way minimizes the intelligence required on the phone by having the phone transmit the data to the Internet-based sever over the cellular connection, as usual, with the server applying the activity recognition model and transmitting the results back to the phone. In one variant, the phone will send the raw accelerometer data and in a second variant the phone will perform the data transformation step and only transmit the data when an example is generated. The second method involves implementing the activity recognition model directly on the cell phone. Given the computational power of these devices, this is certainly a feasible option. One key advantage of this method is that it removes the need for a server, which makes the solution perfectly scalable, and ensures the user's privacy, since the sensor data is kept locally on the device.

The work described in this paper is part of a larger effort to mine sensor data from wireless devices. We plan to continue our WISDM project, applying the accelerometer data to other tasks besides activity recognition and collecting and mining other sensor data, especially GPS data. We believe that mobile sensor data provides tremendous opportunities for data mining and we intend to leverage our Android-based data collection/data mining platform to the fullest extent possible.

6. REFERENCES

[1] Anderson, I., Maitland, J., Sherwood, S., Barkhuus, L., Chalmers, M., Hall, M., Brown, B., and Muller, H. 2007. Shakra: Tracking and sharing daily activity levels with un-augmented mobile phones. In *Mobile Networks and Applications*. 12(2-3).

[2] Apple iPhone and Apple iPod Touch. 2009. Apple Inc. www.apple.com.

[3] Bao, L. and Intille, S. 2004. Activity Recognition from User-Annotated Acceleration Data. *Lecture Notes Computer Science 3001*, 1-17.

[4] Brezmes, T., Gorricho, J.L., and Cotrina, J. 2009. Activity Recognition from accelerometer data on mobile phones. In *IWANN '09: Proceedings of the 10th International Work-Conference on Artificial Neural Networks*, 796-799.

[5] Cho, Y., Nam, Y., Choi, Y-J., and Cho, W-D. 2008. Smart-Buckle: human activity recognition using a 3-axis accelerometer and a wearable camera. In *HealthNet*.

[6] Choudhury, T., Consolvo, S., Harrison, B., LaMarca, A., LeGrand, L., Rahimi, A., Rea, A., Borriello, G., Hemingway, B., Klasnja, P., Koscher, K., Landay, J., Lester, J., Wyatt, D., and Haehnel, D. 2008. The mobile sensing platform: An embedded activity recognition system. In *IEEE Pervasive Computing*, 7(2), 32-41.

[7] Gyorbiro, N., Fabian, A., and Homanyi, G. 2008. An activity recognition system for mobile phones. In *Mobile Networks and Applications*, 14(1), 82-91.

[8] Inooka, H., Ohtaki, Y. Hayasaka, H. Suzuki, A., and Nagatomi, R. 2006. Development of advanced portable device for daily physical assessment. In *SICE-ICASE, International Joint Conference*, 5878-5881.

[9] Krishnan, N., Colbry, D., Juillard, C., and Panchanathan, S. 2008. Real time human activity recognition using tri-Axial accelerometers. In *Sensors, Signals and Information Processing Workshop*.

[10] Krishnan, N. and Panchanathan, S. 2008. Analysis of Low Resolution Accelerometer Data for Continuous Human Activity Recognition. In *IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP 2008)*. Pages 3337-3340.

[11] Lester, J., Choudhury, T. and Borriello, G. 2006. A practical approach to recognizing physical activities. *Lecture Notes in Computer Science: Pervasive Computing*, 1–16.

[12] Mathie, M., Celler B., Lovell N., and Coster A. 2004. Classification of basic daily movements using a triaxial accelerometer. In *Medical & Biological Engineering and Computing*, 42.

[13] Maurer, U., Smailagic, A., Siewiorek, D., & Deisher, M. 2006. Activity recognition and monitoring using multiple sensors on different body positions. In *IEEE Proceedings on the International Workshop on Wearable and Implantable Sensor Networks*, 3(5).

[14] Miluzzo, E., Lane, N., Fodor, K., Peterson, R., Lu, H., Musolesi, M., Eisenman, S., Zheng, X. and Campbell, A. 2008. Sensing meets mobile social networks: The design, implementation and evaluation of the CenceMe application. In *The 6th ACM Conference on Embedded Networked Sensor Systems*, 337-350.

[15] Ravi, N., Dandekar, N. 2005. Activity recognition from accelerometer data. In *Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence*.

[16] Tapia, E.M., Intille, S.S. et al. 2007. Real-Time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In *Proceedings of the 2007 11th IEEE International Symposium on Wearable Computers*, 1-4.

[17] Unwired View.com. 2009. Google wants to make your Android phone much smarter with accelerometer and other sensors. Stasys Bielinis.http://www.unwiredview.com/2009/05/21/google-wants-to-make-your-android-phone-much-smarter-with-accelerometer-and-other-sensors/

[18] Weiss, G. M., and Hirsh, H. 1998. Learning to predict rare events in event sequences, In *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining*, AAAI Press, Menlo Park, CA, 359-363.

[19] WISDM (Wireless Sensor Data Mining) Project. Fordham University, Department of Computer and Information Science, http://storm.cis.fordham.edu/~gweiss/ wisdm/

[20] Witten, I. H. and Frank, E. Data Mining: Practical Machine Learning Tools and Techniques, 2nd ed. Morgan Kaufmann, June 2005.

[21] Yang, J. 2009. Toward physical activity diary: Motion recognition using simple acceleration features with mobile phones, In *First International Workshop on Interactive Multimedia for Consumer Electronics* at ACM Multimedia.

[22] Long, X., Yin, B., and Aarts, R.M. 2009. Single accelerometer-based daily physical activity classification. *In 31st Annual International Conference of the IEEE EMBS*, 6107-6110.

[23] Mannini, A. and Sabatini A.M. 2010. Machine learning methods for classifying human physical activity from on-body accelerometers. In *Sensors 2010*, 10, 1154-1175.

[24] Lee, M., Kim, J., Kim, K., Lee, I., Jee, S.H., and Yoo, S.K. 2009. Physical activity recognition using a single tri-axis accelerometer. In *Proceedings of the World Congress on Engineering and Computer Science 2009*, 1.

[25] Lee, S.-W. and Mase, K. 2002. Activity and location recognition using wearable sensors. In *IEEE Pervasive Computing*, 1(3):24–32.

[26] Subramanya, A., Raj, A., Bilmes, J., and Fox, D. 2006. Recognizing activities and spatial context using wearable sensors. In *Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence*.

[27] Parkka, J., Ermes, M., Korpipaa, P., Mantyjarvi, J., Peltola, J., and Korhonen, I. 2006. Activity classification using realistic data from wearable sensors. In *IEEE Transactions on Information Technology in Biomedicine*, 10(1), 119-128.

[28] Foerster F. and Fahrenberg J. 2000. Motion pattern and posture: correctly assessed by calibrated accelerometers. In *Behavior Research Methods, Instruments, & Computers*, 32(3), 450–7.