AnnoTainted: Automating Physical Activity Ground Truth Collection Using Smartphones

Rahul Majethia
Dept. of Mathematics
Shiv Nadar University, India
rahul.majethia@snu.edu.in

Akshit Singhal
Dept. of Computer Science
University of Texas, Arlington
akshit.singhal2@uta.edu

Lakshmi Manasa K
Dept. of Computer Science
Shiv Nadar University, India
lm160@snu.edu.in

Kunchay Sahiti
Dept. of Computer Science
Shiv Nadar University, India
ks981@snu.edu.in

Shubhangi Kishore
Dept. of Computer Science
Shiv Nadar University, India
sk448@snu.edu.in

Vijay Nandwani
Dept. of Computer Science
Shiv Nadar University, India
vn341@snu.edu.in

ABSTRACT

In this work, we provide motivation for a zero-effort crowd-sensing task: auto-annotated ground truth collection for physical activity recognition. Data obtained through Smartphones for classification of human activities is prone to discrepancies, which reiterates the need for better and larger activity datasets. Artificial data generation algorithms fail to efficiently generate quality instances for minority data. In the proposed model, crowd-sourced sensor data is classified by a robust classifier built by researchers ground up. We nominate a Generic Classifier with $\geq 95\%$ accuracy for this purpose. Data collection and distribution models which ensure that the crowd client receives non-skewed, quality data from locations with higher degree of activity occurrence are elucidated upon. Also integrated within our proposed model are Location-Specific Classifiers, which can be utilized by developers to optimize on location-specific tasks. Effective validation of classified activities using diverse sensor data streams improves the proposed classifier systems and boosts ground-truth accuracy.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

Keywords

Smartphone Sensing, Human Activities, Activity Classification, Ground Truth, Annotation, Classification Complexity

1. INTRODUCTION

The *de-facto* methodology for collecting accurate activity data for any micro- or complex-activity recognition based application has been data collection directly by the researcher or developer whose domain interest lies in the data. Accurate human subject annotated datasets are difficult to come by, and are seldom reliable. Also, when activity datasets are imbalanced or lack adequate representation of each class label, one is forced to resort to artificial data generation algorithms like SMOTE [6]. However, the incapability of these algorithms to restrain from creating feature overlaps between classes drives us to think of ways to generate real-life data for under-represented activities. Furthermore, insitu data collection on certain subjects creates a personification bias in the training data, thus making it erroneous for out-of-sample classification or validation.

In parallel, recent works have proposed a hybrid of crowd-working platforms like Amazon Mechanical Turk[1] and CrowdFlower[2] along with the Sensing-as-a-Service paradigm[18] to offer zero-effort opportunities for Smartphone users to make opportunistically collected sensor-data available. This proposition, bartered in exchange for monetary incentives, could be a popular and scalable way for researchers and developers to acquire well-indexed and context-specific datasets as and when required.

In AnnoTainted, we provide a concise but clear description of such a crowd-sensing task. The issue with collecting correctly annotated ground truth for human activities is two-fold. *First*, lack of motivation as well as insufficiency of validation methods in this area creates an overhead of incorrectly annotated patches. *Second*, even with induced integrity of the collector, deeper insights into the activity streams are missed due to inadequate indexing of such datasets over time. We wish to exploit the ubiquity of the accelerometer in Smartphones to extract suitable features for activity recognition. We propose the use of a robust classifier, with balanced class representation and $\geq 95\%$ cross-validation accuracy, to be used as the benchmark for auto-annotation of data from the crowd-worker. We define the term crowd-worker as a generic human data collector (paid or unpaid). Through AnnoTainted, we aim to make the following key contributions: **(a)** we propose an indexing scheme for activity-data that is based on orthodromic geo-structures (as explained in subsection 3.1) which will, **(b)** provide a probabilistically accurate activity-data inventory

[1] https://www.mturk.com/mturk/welcome
[2] http://www.crowdflower.com/

WPA'16, June 25-30 2016, Singapore, Singapore

ISBN 978-1-4503-4328-2/16/06…$15.00

DOI: http://dx.doi.org/10.1145/2935651.2935653

for custom research applications in human activity recognition, and **(c)** using location-specific data, reduce candidate class labels and hence the classification complexity budget for Smartphones.

We now highlight two sample scenarios where AnnoTainted can be applied:

1. Harry visits a park with his walking group, all of whom have installed AnnoTainted's proposed crowdworker system on their Smartphones. Using our smart data assemblage system described in subsection 4.1, the system would already have classified the park as a source of recurring data for activities such as walking and jogging. When a developer (crowd-client) requests for walking data, Harry and his group would be prime crowd-workers, and the park a prime location, to source the data from. This ensures that the developer has access to the required amount of quality data for the requested activities without the need for artificial data generation.

2. Consider a situation where a City Council wishes to construct pedestrian-only lanes to facilitate the safety and convenience of citizens. This creates a requirement for activity information specific to that city. AnnoTainted's proposed geographical indexing scheme defines *tiles* (see subsection 3.1) which act as constituent elements of the city area. The Council thereby uses a collection of AnnoTainted's Location-Specific Classifiers (described in subsection 5.2) which provide intuition with respect to areas or lanes frequented by pedestrians, i.e., tiles with high occurrence of walking or jogging. This will in turn enable the Council to plan a layout accordingly.

The following pages describe the proposed AnnoTainted framework in detail. The remainder of the paper is organized as follows. In subsection 3.1, we elucidate strategies of achieving location specificity of activities. In subsection 3.2, we define the structure and composition of the robust classifier that helps *co-train* the place specific classifier. The applications of the two key aspects as above are made clearer in sections § 4 and § 5, with explicit emphasis on being able to correctly classify micro-activities and analyze higher-activities on a large scale.

2. RELATED WORK

2.1 Unmonitored Activity Classification

Activity recognition using inbuilt Smartphone sensors has been a steadily growing research area in recent years. The most preferred sensor data streams are from the accelerometer, as it is a low-cost, energy-efficient sensor [15]. Roy et al. in [16] list and describe the features extracted from raw accelerometer data to recognize Activities of Daily Living (ADLs). We use a super-set of the same to classify unmonitored activities. However, the foremost difficulty that arises in case of activity classification outside controlled settings is obtaining well-annotated data streams. Several novel means have been proposed to overcome this obstacle. For instance, [20] illustrates the use of prior knowledge models for activity recognition using data mined from the location-driven social network Foursquare[3] to gain contextual indications. Bhattacharya et al. in [3] suggest an unsupervised learning approach that utilizes unlabelled data for human activity recognition. In contrast, most supervised learning based techniques rely on the end user being asked to annotate the data himself, as is illustrated in [2]. To minimize the error due to human intervention in such approaches, we suggest a labelling method based on a high-accuracy classifier to aid unmonitored activity data collection from Smartphone sensors.

2.2 Need for Quality Training Data

The most prevalent challenge in obtaining quality training data is class imbalance and ground truth insufficiency, which is commonly overcome using *Synthetic Minority Oversampling TEchnique (SMOTE)*[6]. SMOTE generates synthetic samples along the line between the minority class and its nearest neighbor. Further improvements in accuracy have been brought about using Borderline-SMOTE[8], which only over-samples the borderline instances (since they are more likely to be misclassified). Safe-Level-SMOTE[5] assigns weights (safe-levels) to instances and synthesizes minority instances around larger safe-levels. He et al. in [9] propose an adaptive synthetic sampling approach to reduce class imbalance as well as shift classification decision boundaries. However, all these methods manufacture artificial data which is prone to variation and generalization, as shown by De Souza in [7]. AnnoTainted aims to eliminate the need for such simulated data by crowdsourcing real, accurate sensor streams.

2.3 Incentivized Crowdsensing

Various approaches have been suggested to crowdsource sensor data to cater to the needs of researchers and developers. For instance, Lane et al. in [13] propose 'piggyback crowdsensing', where data is collected opportunistically based on device usage patterns. However, a majority of GPS data required to design the mobility model described by the authors is sourced from GPS samples requested for by other applications. This poses a problem in the case of a large number of Android devices since the Android Application Sandbox isolates individual app data and permissions[17]. This presents a case for an incentive-driven crowdsourcing model, where the crowd-worker gives explicit permissions to the crowd data collection application. Moreover, crowd-working platforms like Amazon Mechanical Turk or CrowdFlower are gaining popularity as supplementary sources of income. However, they are platform- and device-independent and do not utilize the resource availability and mobility of the crowd-worker's Smartphone. To take advantage of the same, Sheng et al. in [18] propose an incentive-driven paradigm called Sensing-as-a-service to collect sensor data from Smartphones via a cloud-computing system. AnnoTainted proposes a similar incentivized model, though without human intervention.

3. SYSTEM FRAMEWORK

The following section describes a set of step-by-step strategies and components to develop a baseline activity regulation system. It discusses in detail each of the aspects of human activity recognition that we keep in mind while designing AnnoTainted. As mentioned in § 1, AnnoTainted is proposed to be a completely intervention-free system. Therefore, it must evade the pitfalls of standardization incon-

[3] https://foursquare.com/

sistencies that plague the area of tapping such data from Smartphones. In order to do so, obtaining contextual data and the construction of robust activity classifiers remain our primary aims.

3.1 Achieving Context

In AnnoTainted, the key requirement is to be able to achieve fine-grained context regarding activity patterns in distinct location structures. We bring about an amalgamation of the same using the following *two* key context achieving concepts.

3.1.1 Area Interpretation

The area under observation for collection of data instances is divided into atomic units to associate detailed descriptions about physical activities identified in a said unit. We define an orthodromic area unit called tile, which is a 3D structure of variable dimensions. The orthodromic character of the tile corresponds to the shortest distance between two points along an arc. This allows us to take into account the curvature and contours of the earth while determining the dimensions. The dimensions of the tile are determined by fitting them in a manner such that the periphery of an area of interest is defined by the tiles. Figure 1 is a representation of an area of interest divided into tiles.

3.1.2 User Location

The placement of the crowd-worker within a defined tile is achieved through GPS data. GPS, along with providing the location details, associates with it a confidence value. This is a measure of the range within which the provided location is accurate. As long as the confidence value falls within the dimensions of the tile, system accuracy remains consistent as the location of the crowd-worker is attributed with the appropriate tile. As a remark, it should be noted that Android provides accuracy only with 68% confidence. In statistical terms, it is assumed that location has a randomly distributed random error, so the 68% confidence circle represents one standard deviation. However, such a simple distribution might not be followed in real-life applications.

3.2 Robust Generic Classifier

3.2.1 Activity Ensemble

The micro-activity (MA) set chosen for AnnoTainted consists of the following primary physical activities: stationary, walking, jogging, commuting (via motorized transport), ascending and descending stairs. These activities are the fundamental human physical movements classified through accelerometer data in existing research such as [12].The recurrence of these activities in daily human routine and their constituent repetitive motions enable easy recognition. Higher activities (HAs) can be defined as those obtained through permutations of the micro-activity ensemble. Such HAs form an extensive set and can generally be detected from MAs. Hence our principal focus in this work will be acquiring data for and classifying aforementioned micro-activities only.

3.2.2 Frequency and Sample sizes

The sampling rate, sampling stability and noise can give us the quality of the acceleration sensor in Smartphones. AnnoTainted audits the in-built accelerometer sampling frequency in the crowd-worker's Smartphone and gives data collected from the device a priority number accordingly. A suitable minimum sampling frequency (f_m) is chosen and devices with sampling frequency less than this threshold are invalidated for data collection as the results produced by such devices are not considered to be accurate enough for activity recognition. Higher sampling rates generally lead to more accurate values due to a larger number of samples being accessible in the time window. f_m is set according to recognition accuracy obtained through testing at various sampling rates. The recognition accuracy is defined as the number of accurately classified features for the defined activity ensemble. The f_m in our system is set to 32Hz, a sampling rate that is sufficient for potential recognition of body movements using accelerometer data, and is also supported by prior research in [14].

3.2.3 Placement and Orientation

Smartphone placement and orientation during sensor data collection alter accelerometer and gyroscope data, which in turn affect the classification of physical activities. This enforces the need for error correction and higher accuracy rates in the measurement of the aforementioned sensor data. It was established in [11] that consideration of orientation and rotational variations achieved accuracies upto 85%. In order to achieve higher accuracy in the sensor data obtained through AnnoTainted, we use a rotation based approach as described in [19], with angular velocity and rotation radius as the elementary features for classification.

3.2.4 Feature Sets

Periodic patterns for jogging, walking and movement on stairs are characterized by parameters such as time periods between observed peaks and magnitudes of acceleration. Hence the feature set chosen for classification is a combination of the feature sets suggested in [16], where 3-axis accelerometer and gyroscope data was used to compute a 30-dimensional feature vector over successive time frames, and [12], where peak acceleration values were primarily considered to generate 6 basic features and 43 summary features. For commuting, the feature set suggested in [10](specific to transportation mode detection) was used, where peak, frame and segment based features were considered, including breaking periods and stopping rate features.

3.2.5 Balanced Activity Representation

To ensure the model is not biased towards majority labels, the training dataset needs to be re-calibrated if it is skewed and certain activities are under-represented. Since the Generic Classifier is built ground up using activity data collected by researchers, we ensure that equal, correctly labelled instances of each activity are included.

4. QUALITY DATASET GENERATION

In order to best approximate correctly annotated data generated under controlled conditions that is both relevant and consistent, the patterns of activity aggregation have to be determined. In order to establish the most pertinent location, we observe the nature and trends of the data obtained, elaborated further in this section.

4.1 Data Assemblage and Activity Trends

The instances of data collected from a tile are continuously categorized using the generic classifier described in § 3. This subsection describes the probability based quantification of a tile with the likelihood of a particular activity occurring in the said region. This type of activity trend recognition is important in order to effectively classify regions as *hotspots* for collection of data for certain types of activities. Systems such as in [20] perform activity trend identification by utilizing tips obtained through social media. We seek to achieve a similar result through recognition and classification of real time activities. To this end, we implement Exponential Weighted Moving Average (EWMA) to model the probability of a particular activity as a function of the past occurrences of the said activity in a specific geographical domain, hence reducing rapid morbidity in the recorded activity *trend* of a place. The recurrence value (R) of an activity, a_i happening in an tile j, is described as

$$R_{ij} = \alpha * P_{n-1}(x_i = 1) + (1 - \alpha) * P_n(x_i = 1)$$

where x_i is the random discrete variable that describes the occurrence of activity a_i and n is the number of recorded instances until that point of time. α is the degree of weighting decrease, which is indicated by a value between 0 and 1. The value for this smoothing parameter α can be selected on the basis of the importance to be given to current trending activities, the most safest option being proportional weights to all recorded instances. For instance, if the previously calculated recurrence value for walking in a park, P(a_i,T_j), is 0.830, and the probability of data instance collected being true for the said activity is 1, with the data instance being the 68th such collected record,

$$P(a_i, T_j) = 0.985 * 0.830 + 0.014 * 1 = 0.831$$

This calculation, along with presenting the method, goes to show that the continuously updating values of P(a_i,T_j) increase fractionally. The amount by which this increase is quantified (α) is a value that is inversely proportional to the number of records (n). Hence, the probability P(a_i,T_j) becomes less volatile as n increases. Another point to be noted is that as the R value for one activity is updated, the R values of all other activities in that tile are also altered due to the non-occurrence of this activity in the data instance set.

4.2 Optimizing Tile Dimensions

Here, we seek a trade-off between the size of the tile and the amount of data instances that can be collected from it. The smaller the size, the better the approximation of the periphery of the structure. This implies that a minimal volume of the tile exists outside the boundary of the structure. However, as the dimensions of the tile decrease, the assemblage of data instances from the tile that has the potential to model meaningful data is restricted. This presents an optimization problem which can be represented as:

$$\begin{array}{ll} \underset{T}{\text{maximize}} & \dfrac{Area(s_j)}{Area(T_j)} \\ \text{subject to} & f_i(x) \leq b_i, \; i = 1, \ldots, m. \end{array}$$

where $Area(T_j)$ represents the area of the Tile j and $Area(s_j)$ is the amount of the structure that lies within tile T_j. The minimum constraint functions f_i can be one or more of $P(a_i, T_j)$, $N(a_i, T_j)$ etc. (see Algorithm 1 for terminology). Although the variable dimensions allow for a diverse range of structures to be observed, for our preliminary platform we use tiles of fixed dimensions. We integrate a solution to the optimization problem into our dynamic tile system as a part of our future work in this field.

4.3 Data Fetching Algorithms

When a crowd-client requests for annotated data corresponding to a particular activity, AnnoTainted aims to fulfill it while adhering to two inherent quality constraints. Firstly, the data requirement of the crowd-client, in terms of number of instances, should be fulfilled to the best extent possible. Secondly, the accuracy of annotation must be kept as high as possible. To this end, the annotated instances must be probabilistically validated. This optimization problem can be viewed as one that loosely translates to the well known *knapsack problem*. An analogous solution is proposed and described below.

We begin to source the data from the tile that has the highest recurrence value for that activity. This is accomplished using the arguments of the maxima of the set of recurrence values of a specific activity a_i over all the tiles until the data requirement of the crowd-client is fulfilled. Figure 3 is a representation of the most appropriate location (According to Algorithm 1) to source the depicted activities from.

```
Algorithm 1: AnnoTainted - Data Fetching
Data: $\widetilde{A}$, activity subset wanted
Data: $N(\widetilde{a}_i)$, # of instances wanted for activity $\widetilde{a}_i$
Data: $P(\widetilde{a}_i, T_j)$, Probability of $\widetilde{a}_i$ at tile $T_j$
Data: $I(\widetilde{a}_i, T_j)$, Instance Set of $\widetilde{a}_i$ at tile $T_j$
Result: $I_F$ , Final Quality Data-set
Function knapsackActivityData($\widetilde{A}, N$)
    $I_F = \phi$
    /* For all Required Activities                              */
    for $i = 1, 2, \ldots, |\widetilde{A}|$ do
        /* Pick Data from high P($\widetilde{a}_i$)                   */
        sortInDescending($P(\widetilde{a}_i, T)$)
        for $j = 1, 2, \ldots, |\boldsymbol{T}|$ do
            if $\widetilde{I} + I(\widetilde{a}_i, t_j) \leq N(\widetilde{a}_i)$ then
                $\widetilde{I} = \widetilde{I} \cup I(\widetilde{a}_i, t_j)$
            else
                /* $n = |N(\widetilde{a}_i)| - |\widetilde{I}|$                          */
                $\widetilde{I} = \widetilde{I} \cup (I_n \overset{n}{\subset} I(\widetilde{a}_i, t_j))$
                break
        $I_F = I_F \cup \widetilde{I}$
        $\widetilde{I} = \phi$
    return $I_F$
```

4.4 Activity Validation

After physical activity classification, their validation is of significant importance in order to eliminate incorrectly classified activities. This can be accomplished through the application of additional sensor data streams. For a preliminary stage validation, we utilize data streams from the GPS on the crowd-worker's Smartphone to mine useful parameters such as velocity (V_{GPS}), altitude and acceleration(A_{GPS}). The aforementioned data is used to assist in determining *thresholds* or restrictions which can disqualify certain misclassified instances. We use this data in addition to the

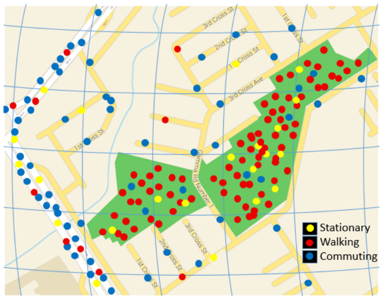

Figure 1: Preliminary stage Tile Definition

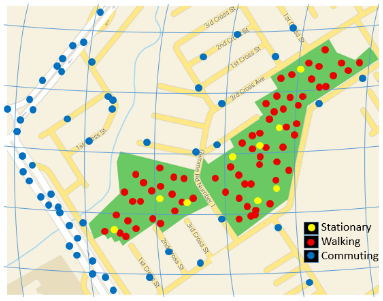

Figure 2: After Activity Validation

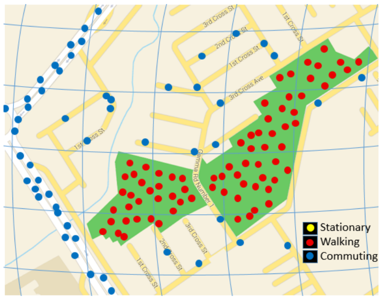

Figure 3: Location based activity preference

classified activity to generate key-value pairs where the activity is considered appropriately classified if it conforms to the constraints summarized in Table 1. Figure 2 is a representation of the activities after validation.

Table 1: Activity Validation Constraints

Activity	Threshold
Walking	$V_{GPS} < 5.7$ kmph
Stationary	$V_{GPS} \approx 0\ km/h$ & $A_{GPS} \approx 0\ m/s^2$
Commuting	$A_{GPS} < 4\ m/s^2$

5. LOCATION-SPECIFIC DATA

In the previous section, we discussed the plausibility and methodology of collecting high accuracy ground-truth for trending activities from multiple structural domains. In this section, we discuss the motivation behind using the data collected within a Tile *only* and using the same for any location-specific application within the same geographical structure. We elucidate the trade-offs of this approach and present application scenarios of the same.

5.1 Raw Activity Stream

The raw activity data stream of the form $\{(data_1, label_1), (data_2, label_2), ... \}$ is obtained from every Tile T_j. We propose an open availability of the data stream so that developers can use the same to optimize the wide range of location-specific classification based applications. Keeping in mind the geographical layout of the corresponding Tiles, and by analyzing the micro-activity stream, the developer will be able to gain deeper *complex* activity (hiking etc.) insights, as compared to data with no context.

Solving another problem, location-sensitive data could possibly result in representation of only a subset of activities in any given tile T_j. As the implementation of a multi-class classifier calls for $l(l-1)/2$ binary classifiers, each trained with data from a distinct pair of classes, a reduced number of class-labels is only a boon to any Smartphone-sensing application, as the classification can now take place on devices with a low-energy budget.

With an increase in the number of class labels, not only do the number of decision hyperplanes increase, there is also a fall in accuracy due to wrongly classified points when the kernel functions are simple. With more complex functions, it is obvious that per-classification time is compounded. Numerous approaches to counter the complexity have been proposed, such as those detailed in [1],[4]. However, we suggest eliminating the cause of the problem, to some extent, by reducing the number of class labels to $(l-k)$, as described in the next subsection.

5.2 Location-Specific Classifier

We propose building a Location-specific Classifier (LC) for a Tile T_j with possible class labels $a_1, ...a_k \subseteq \mathbf{A}$. Being a k-nomial classifier rather than an l-nomial one, where $l = |\mathbf{A}|$ and $k \leq l$, the LC is faster and more cost-effective computationally. We use supervised learning algorithms to train the LC using the accelerometer feature vectors that have been labelled by the GC. In doing so, we use only the instances collected from a particular geographical Tile T_j. As more instances emerge from T_j, we apply an incremental learning approach to include these in the training set of the LC. To do so, we define an *epoch* - a set I_E of instances with a fixed cardinality c. After every epoch of feature vectors collected from the tile T_j (after every set of c instances) being classified by the GC, the LC is updated taking into account this labeled data.

5.2.1 *Loss of Accuracy*

The trade-off in using a more specific $(l-k)$ class classifier rather than the generic l-class one is the loss of accuracy. This is because the LC is trained only on the basis of instances that have been labelled by the GC until the end of the last epoch. Any instance of a new activity is therefore not present in the training set of the LC, and is hence treated as an anomaly. To overcome this challenge, the epoch after which the LC is remodeled has to be sufficiently short. However, he usefulness of the LC is limited by its poor accuracy in comparison to the GC. In order to reduce computational latency and still maintain a threshold level of accuracy (determined by the purpose of usage), we suggest a mechanism as described below.

5.2.2 *Classifier Confidence Correction*

Every time a user arrives in a region of interest, location-specific classifiers are used in either conjunction or alternation with the generic validating classifier, to schedule a sequence of classification of MA$_s$. In real time, choosing the classifier should be done by a weighted random scheduling algorithm. For any given person, for a confidence threshold $\epsilon \in \mathbf{R}[0,1]$, we use the LC with probability ϵ and *both* LC and GC with probability (1 - ϵ). Every time both the classifiers are used, a match (or mismatch) of the classified label updates the confidence value of the LC. We propose an epoch to end when the confidence of the classifier has fallen beyond a particular threshold and cannot be sustained any longer without updating. The application developer entity is liable to set the required accuracy in order to schedule the alternation of the LC and the GC accordingly.

6. CONCLUSION AND FUTURE WORK

In this paper, we discuss a methodology for crowdsourcing classifier-annotated sensor data for physical activity recognition. We propose a geographical matrix of Tiles to segregate the collected data and learn the degree of recurrence of each activity in a particular domain. Based on this scheme of indexing, we are able to provide quality datasets for custom application developers and researchers. Further, we use this location-specific data to reduce classification complexity in domain-restricted applications.

As future work, we plan to impose temporal domain restrictions on data collection in addition to the spatial restrictions described in subsection 4.3 by sourcing activity data from periods with highest likelihood of the activity being performed. Also, we aim to devise a payment model for the crowd-worker, keeping in mind the hardware capabilities of the device and the degree of mobility of the user. In addition to GPS data, we also plan to source data from alternative sensors to design contextually relevant restrictions.

7. ACKNOWLEDGEMENTS

We would like to thank our colleagues, whose help in aggregating activity data for out prototype system was invaluable, and Dr. Krishnan Rajkumar for his advice, support and comments which have greatly helped furnish this work better. We would also like to acknowledge and express our gratitude to the anonymous reviewers for their constructive comments and critical feedback, most of which we have been able to incorporate in this version of our work.

8. REFERENCES

[1] E. L. Allwein, R. E. Schapire, and Y. Singer. Reducing multiclass to binary: A unifying approach for margin classifiers. *The Journal of Machine Learning Research*, 1:113–141, 2001.

[2] L. Bao and S. S. Intille. Activity recognition from user-annotated acceleration data. In *Pervasive computing*, pages 1–17. Springer, 2004.

[3] S. Bhattacharya, P. Nurmi, N. Hammerla, and T. Plötz. Using unlabeled data in a sparse-coding framework for human activity recognition. *Pervasive and Mobile Computing*, 15:242–262, 2014.

[4] E. J. Bredensteiner and K. P. Bennett. Multicategory classification by support vector machines. In *Computational Optimization*, pages 53–79. Springer, 1999.

[5] C. Bunkhumpornpat, K. Sinapiromsaran, and C. Lursinsap. Safe-level-smote: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem. In *Advances in knowledge discovery and data mining*, pages 475–482. Springer, 2009.

[6] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: Synthetic minority over-sampling technique. *J. Artif. Int. Res.*, 16(1):321–357, June 2002.

[7] C. De Souza. Classification of imbalanced classes.

[8] H. Han, W.-Y. Wang, and B.-H. Mao. Borderline-smote: a new over-sampling method in imbalanced data sets learning. In *Advances in intelligent computing*, pages 878–887. Springer, 2005.

[9] H. He, Y. Bai, E. A. Garcia, and S. Li. Adasyn: Adaptive synthetic sampling approach for imbalanced learning. In *Neural Networks, 2008. IJCNN 2008.(IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*, pages 1322–1328. IEEE, 2008.

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

[11] O. D. Incel. Analysis of movement, orientation and rotation-based sensing for phone placement recognition. *Sensors*, 15(10):25474–25506, 2015.

[12] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, 12(2):74–82, 2011.

[13] N. D. Lane, Y. Chon, L. Zhou, Y. Zhang, F. Li, D. Kim, G. Ding, F. Zhao, and H. Cha. Piggyback crowdsensing (pcs): energy efficient crowdsourcing of mobile sensor data by exploiting smartphone app opportunities. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, page 7. ACM, 2013.

[14] S. L. Lau and K. David. Movement recognition using the accelerometer in smartphones. In *Future Network and Mobile Summit, 2010*, pages 1–9. IEEE, 2010.

[15] A. Rai, Z. Yan, D. Chakraborty, T. K. Wijaya, and K. Aberer. Mining complex activities in the wild via a single smartphone accelerometer. In *Proceedings of the Sixth International Workshop on Knowledge Discovery from Sensor Data*, pages 43–51. ACM, 2012.

[16] N. Roy, A. Misra, and D. J. Cook. Ambient and smartphone sensor assisted ADL recognition in multi-inhabitant smart environments. *J. Ambient Intelligence and Humanized Computing*, 7(1):1–19, 2016.

[17] B. P. Sarma, N. Li, C. Gates, R. Potharaju, C. Nita-Rotaru, and I. Molloy. Android permissions: a perspective combining risks and benefits. In *Proceedings of the 17th ACM symposium on Access Control Models and Technologies*, pages 13–22. ACM, 2012.

[18] X. Sheng, J. Tang, X. Xiao, and G. Xue. Sensing as a service: Challenges, solutions and future directions. *Sensors Journal, IEEE*, 13(10):3733–3741, 2013.

[19] Y. Shi, Y. Shi, and J. Liu. A rotation based method for detecting on-body positions of mobile devices. In *Proceedings of the 13th international conference on Ubiquitous computing*, pages 559–560. ACM, 2011.

[20] Z. Zhu, U. Blanke, A. Calatroni, and G. Tröster. Prior knowledge of human activities from social data. In *Proceedings of the 2013 International Symposium on Wearable Computers*, pages 141–142. ACM, 2013.