Poster: am i indoor or outdoor?

Citation for published version:
Radu, V, Katsikouli, P, Sarkar, R & Marina, MK 2014, Poster: am i indoor or outdoor? in *The 20th Annual International Conference on Mobile Computing and Networking, MobiCom'14, Maui, HI, USA, September 7-11, 2014.* ACM, pp. 401-404. https://doi.org/10.1145/2639108.2642916

Digital Object Identifier (DOI):
10.1145/2639108.2642916

Published In:
The 20th Annual International Conference on Mobile Computing and Networking, MobiCom'14, Maui, HI, USA, September 7-11, 2014

Poster: Am I Indoor or Outdoor?

Valentin Radu
valentin.radu@ed.ac.uk

Panagiota Katsikouli
p.katsikouli@sms.ed.ac.uk

Rik Sarkar
rsarkar@inf.ed.ac.uk

Mahesh K. Marina
mahesh@ed.ac.uk

The University of Edinburgh, UK

ABSTRACT

The environmental context of a mobile device determines where/how it is used, which can be exploited for efficient operation and better usability. In this work we describe a general method using only the lightweight sensors on a smartphone to detect if a device is indoor or outdoor. Using semi-supervised machine learning techniques, our method automatically learns characteristics of new environments and devices, thereby achieves detection accuracy of over 90% even in unfamiliar circumstances. Therefore, it easily outperforms existing indoor-outdoor detection techniques based on static algorithms, or relying on energy hungry and unreliable GPS.

Categories and Subject Descriptors

C.3 [**Special-Purpose and Application-Based Systems**]: Signal processing systems; D.4.8 [**Performance**]: Modeling and prediction

General Terms

Design, Experimentation, Performance

Keywords

context detection; indoor-outdoor; smartphone sensing; machine learning

1. INTRODUCTION

With rapid adoption of smartphones, context sensing/detection is becoming increasingly important to enable new and sophisticated context-aware mobile apps. There are various forms of context concerning a mobile user including location, environment, time and activity. While time is straightforward to identify with a smartphone, other aspects of context are harder to determine and require more advanced integration of several inbuilt sensors on modern smartphones.

The focus of this work is on detecting whether a user is indoor or outdoor, an aspect of environmental context, which we refer to as the *Indoor-Outdoor (IO) Detection* problem. This user context information is important for a variety of applications including personalization applications (change volume, screen brightness, application shortcuts), power saving (turn off GPS indoors, turn off WiFi radio outdoors) and reminders. Currently, there are two existing techniques for IO detection. One is based on GPS and uses the drop in confidence or inability to obtain a fix as a cue to infer that the user is indoors (e.g., [3, 4]). The other called IODetector [5] combines estimations from cell signal, light and magnetic intensity based features.

In this work, we first experimentally examine the effectiveness of these two techniques in different real-world settings in detail, and find that neither of them provides satisfactory results. GPS is unreliable because it is sometimes possible to get GPS fix while indoors and not get it in some outdoor locations. More crucially, GPS is among the most energy hungry smartphone sensors. On the other hand, the accuracy with IODetector is quite poor because it does IO detection using fixed thresholds for sensor features that are not appropriate across different environments.

Motivated by the above observations, we propose a new approach to IO detection that is based on *semi-supervised learning*. Underlying our proposal is a model that continuously adapts as the user visits new environments, with characteristics distinct from those seen previously, by obtaining new training data on the fly without user involvement. Like IODetector, it still relies only on low power sensors and can be fully implemented on the phone. As a result, it overcomes the limitations of existing techniques. Specifically, our method can achieve very high accuracy (greater than 90%) across diverse and unseen environments while being energy-efficient.

2. PREVIOUS SOLUTIONS

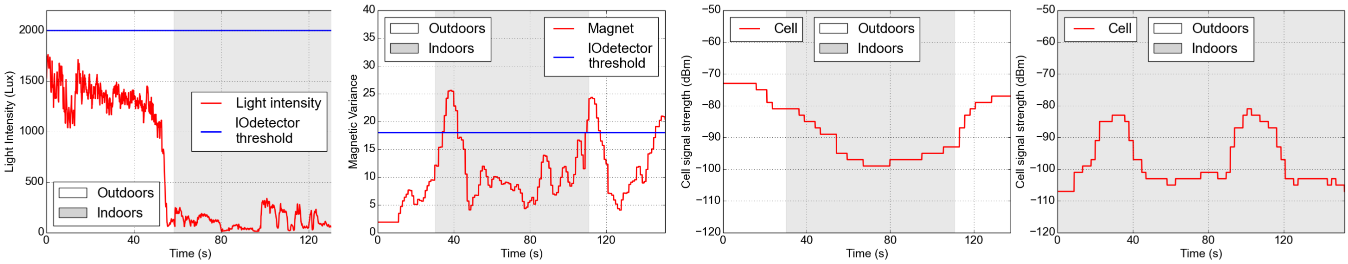
2.1 GPS based Indoor-Outdoor detection

GPS signals are usually available outdoors where the sky is directly visible, and are often weak or unavailable indoors when the sky is obscured by ceiling and walls. Thus, the estimated accuracy of GPS localization can be used to detect if a user is indoors [3].

The primary drawback of GPS is that it is among the most energy hungry sensors on the phone. We observed that with a power consumption of 370mW, the GPS uses one order of magnitude more power than the microphone, the magnetic and the light sensors, and up to two orders of magnitude more power than the accelerometer and the battery thermometer, for continuous sensing.

Moreover, GPS can sometimes get a satellite fix indoors (e.g., when the user is close to a door or window), which reduces its reliability as an indoor-outdoor classifier. In our experiments we found that indoor/outdoor state can be correctly determined with an accuracy around 70-80%, even with a carefully chosen threshold on GPS localization accuracy as shown later.

MobiCom'14, September 7-11, 2014, Maui, Hawaii, USA.
ACM 978-1-4503-2783-1/14/09.
http://dx.doi.org/10.1145/2639108.2642916.



(a) Light detector (b) Magnetic detector (c) Cell signal at transition (d) Cell signal indoor

Figure 1: Illustration of cases where IODetector components (light, magnetic and cell) fail to correctly distinguish between indoors and outdoors. (a) The light intensity is always below the threshold. (b) The magnetic variance is low indoors. (c) The cell signal strength decreases slowly for cell component to detect transition. (d) The cell component detects transitions indoors while moving between rooms due to rapid signal variations. Background shaded (non-shaded) portions in the figures indicate actual ground truth, indoor (outdoor).

2.2 IODetector

IODetector [5] is a recent work that uses low power sensors available on smartphones (light, cell radio and magnetic field sensors) to determine indoor, outdoor and semi-outdoor states. Each of these three sensors works independently to estimate a state and a confidence value. The final state is determined after adding all the confidence values.

Fundamentally, IODetector is built based on some experimental observations: (1) In daytime, in outdoors, light intensity is typically above 2000 Lux; (2) When the user's context changes from outdoors to indoors, the cell signal strength drops rapidly due to attenuation from walls and ceilings; and (3) Magnetic field sensed by the phone tends to change rapidly when the user is moving indoors where there are possibly many appliances, electric currents and metallic objects nearby, compared to outdoors.

These observations may hold in some cases, but we find that they do not hold always and across different environments. Because IODetector has hard-coded thresholds for all the three components (light, cell and magnetic), the system is invariant to changes in relevant factors like environments, weather conditions, seasons, latitude and devices, which ultimately hurts the accuracy of the IO detection.

Figure 1 presents some cases where the IODetector components fail to detect the right IO state because of the non-adaptive nature of their respective thresholds. In figure 1(a), even though the light intensity changes drastically at the transition from outdoor to indoor, the component fails to distinguish between the two states because the light level is always below the threshold, even outdoors. In figure 1(b), the magnetic variance is below the threshold for most of the time, both indoors and outdoors. The cell signal decreases when moving from outdoor to indoors in figure 1(c), but the speed with which this happens is too slow for the component to detect the transition. On the other hand, in figure 1(d) the cell component erroneously detects IO transitions when moving between the rooms of the same building.

3. ROBUST INDOOR-OUTDOOR DETECTION WITH SEMI-SUPERVISED LEARNING

In this section, we present a novel, adaptive IO detection approach that overcomes the problems due to static thresholds as with IODetector, while at the same time avoids the use of energy-hungry sensors like GPS and Wi-Fi. Moreover, unlike IODetector, we are interested mainly in the basic states, which are indoor and outdoor, since these are the ones most relevant to context-aware applications. A semi-outdoor state considered in IODetector system is difficult for many applications to interpret since the environment characteristics for this state are unpredictable. Indoor/outdoor transitions are objectively defined, by crossing a threshold such as a door, but the determination of a state to be semi-outdoor is fairly subjective. This makes it difficult to obtain meaningful ground truths from users to evaluate the reliability of a method using semi-outdoor state.

Specifically, we employ a semi-supervised learning approach, which targets situations where obtaining annotated information with the ground-truth is impractical or expensive to obtain [6]. An additional advantage of semi-supervised learning techniques is that subtle differences between classes, possibly unobservable in a small amount of training data with actual ground truth information (referred henceforth as labeled data), are more likely to be detected and exploited to achieve more reliable classification.

Among the several most commonly used semi-supervised techniques (e.g., self-training, co-training), we experimentally found co-training [1] to be most effective for the IO detection problem (details omitted due to space restrictions). Co-training is a method where two classifiers work in parallel to improve predictions. These classifiers work with different sensor features to gain different perspectives and uncover different patterns. That is, each data point is classified by 2 different classifiers working with different features/attributes and the result with higher confidence is used to retrain and improve both classifiers. The two perspectives underlying the co-training approach based on different feature sets ensure that the training of classifiers is not affected by same biases. The idea behind co-training is shown schematically in Figure 2. See [6, 2, 1] for more details.

Note that supervised learning based classifier approach has the same fundamental limitation as IODetector in that a supervised learning based classifier model trained in one environment may not be accurate in other environments. While collecting labeled data to train in each new environment would make supervised learning approach work but will be impractical. Semi-supervised learning approach essentially overcomes this problem by generating labeled data on the fly without user involvement.

3.1 Data collection

In addition to the features used by IODetector (magnetic variance, cell signal strength, light intensity, proximity sensors and

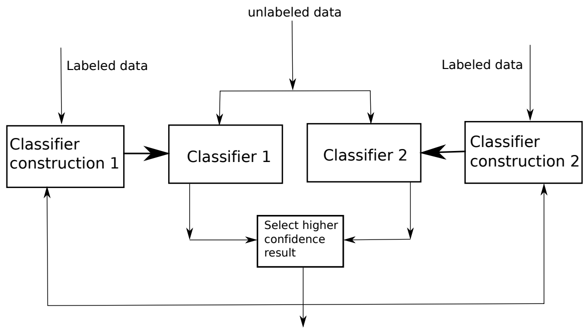


Figure 2: Co-training with 2 classifiers operating with different feature sets. The higher confidence classification for each data is used as the training label to improve classification.

time of day), our method relies on a slightly more expanded set including other light weight features: battery temperature and microphone detected noise amplitude.

To evaluate our semi-supervisied learning (co-training) approach, using a custom Android app, we collected more than 3800 samples of sensor data with a fair distribution between three environments: university campus, city center and residential area. To assess the accuracy of different approaches, the Android app relies on an interface for volunteers who participated in the data collection to manually input (indoor/outdoor) ground-truth information.

3.2 Feature ranking and selection

In building the two classifiers for co-training, we balanced the feature sets in terms of reliability. We ranked the features using different machine learning techniques and then split them into two sets with a fair distribution. Using Naive Bayes analysis of feature importance, we determined the following distribution: Set1: {light intensity, time of the day, proximity value and battery temperature} and Set2: {sound amplitude, cell signal strength, magnetic variance}. Using the SVM Attribute ranking we balanced the two sets into: Set1: {cell signal strength, light intensity, time of day and proximity value} and Set2: {battery temperature, sound amplitude and magnetic variance}. We then evaluated the performance of each of these two feature distributions as part of the classifiers in the co-training approach.

3.3 Evaluation of Co-training

For our co-training method, we chose a small number of labeled instances (300) from one environment (campus) plus 1000 unlabeled instances from the other two environments. For each instance, the system takes the higher confidence classification to be the inferred label for re-training (see Fig. 2). Table 1 presents the accuracy of different configurations of the co-training method with 1200 completely separate instances used for evaluation spread across different environments. The best performance was obtained by a combination of Naive Bayes classifiers using a distribution of features based on the SVM ranking.

In a direct comparison with the previous solutions (IODetector and GPS) and a supervised learning approach (with 300 labeled instances from campus environment for training), our solution using co-training with SVM Attributes distribution has better results by at least 10% (Table 2). These results do not capture other advantages of our approach that include: similar energy consumption as IODetector method; model adaptation without user involvement; and fully implementable on the phone.

Features Distribution	Classifier 1	Classifier 2	Performance (%) Co-training
Naive Bayes based	J48	J48	83.0
	LWL	LWL	78.17
	Naive Bayes	Naive Bayes	91.66
SVM Attribute ranking based	J48	J48	86.67
	LWL	LWL	78.16
	Naive Bayes	Naive Bayes	**93.33**

Table 1: Accuracy with different configurations of co-training method.

Solution	Performance (%)
GPS	75.23
IODetector	35.74
Supervised learning	81.29
Co-training (SVM Attr.)	**93.33**

Table 2: IO detection accuracy for previous solutions (GPS, IODetector), supervised learning on IODetector features and our co-training method.

4. CONCLUSIONS

We have considered the problem of determining whether a user is indoors or outdoors using low power sensors readily available on modern smartphones. We show that existing solutions are too energy hungry or fail to provide accurate results across a range of environments user may typically encounter, due to the use of fixed environment agnostic thresholds in the underlying estimation schemes. To address the fundamental issue of model adaptation on the fly transparent to the user, we employed a novel approach based on semi-supervised learning (specifically, co-training technique). Our adaptive solution not only outperforms existing techniques but also supervised learning based classifier approach. In on-going and future work, we intend to demonstrate the value of our proposed approach in the context of various use cases.

5. ACKNOWLEDGMENTS

We thank Charles Sutton for helpful discussions.

6. REFERENCES

[1] A. Blum and T. Mitchell. Combining Labeled and Unlabeled Data with Co-training. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, COLT' 98, 1998.

[2] O. Chapelle. *Semi-Supervised Learning*. MIT Press, Cambridge, Mass, 2006.

[3] L. Ravindranath, C. Newport, H. Balakrishnan, and S. Madden. Improving Wireless Network Performance Using Sensor Hints. In *Proc. USENIX NSDI*, 2011.

[4] H. Wang, S. Sen, A. Elgohary, M. Farid, M. Youssef, and R. R. Choudhury. UnLoc: No Need to War-Drive: Unsupervised Indoor Localization. In *Proc. MobiSys*, 2012.

[5] P. Zhou, Y. Zheng, Z. Li, M. Li, and G. Shen. IODetector: A Generic Service for Indoor Outdoor Detection. In *Proc. SenSys*, 2012.

[6] X. Zhu. *Semi-Supervised Learning with Graphs*. PhD thesis, Carnegie Mellon University, 2005.