# Transfer Learning for WiFi-based Indoor Localization

**Sinno Jialin Pan, Vincent Wenchen Zheng, Qiang Yang and Derek Hao Hu**
Department of Computer Science and Engineering
Hong Kong University of Science and Technology, Hong Kong
{sinnopan,vincentz,qyang,derekhh}@cse.ust.hk

### Abstract

The WiFi-based indoor localization problem (WILP) aims to detect the location of a client device given the signals received from various access points. WILP is a complex and very important task for many AI and ubiquitous computing applications. A major approach to solving this task is through machine learning, where up-to-date labeled training data are required in a large scale indoor environment. In this paper, we identify WILP as a transfer learning problem, because the WiFi data are highly dependent on contextual changes. We show that WILP can be modeled as a transfer learning problem for regression modeling, where we identify several important cases of knowledge transfer that range from transferring the localization models over time, across space and across client devices. We also share our working experience in WILP and transfer learning research in a realistic problem solving setting, and discuss a data set we have made public for advancing this research.

## Introduction

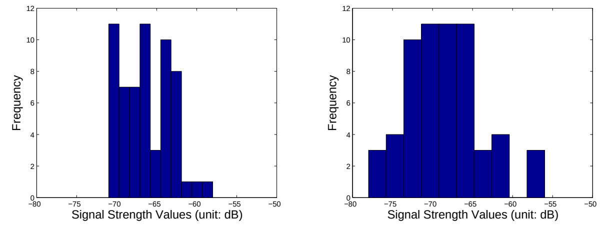
Accurately locating a mobile device in an indoor environment is an important task in many AI and ubiquitous computing applications. Examples can be found ranging from activity recognition, robotics to various user-assisted technologies such as home-based healthcare [1]. With the increasing availability of 802.11b/g WiFi network in various cities and urban centers, indoor localization is increasingly feasible for indoor location detection based on wireless signal strength values (Bahl, Balachandran, & Padmanabhan 2000; Letchner, Fox, & LaMarca 2005; Ferris, Hähnel, & Fox 2006). The WiFi based indoor localization problem (WILP) is also a difficult task because the WiFi data are very noisy and highly dependent upon environment due to multi-path and shadow fading effects in indoor environment. The data distribution is constantly changing depending on various factors, such as human movement, temperature and humidity changes, (Yin, Yang, & Ni 2005). When applying machine-learning based approaches, it is very costly to collect and label the training data in the form of (RSS values, Location Label) pairs in a large scale building, because humans need to take a mobile device and walk through the building to collect the RSS values and mark down the ground truth locations (Ferris, Fox, & Lawrence 2007; Pan *et al.* a). When the signal distribution changes, such processes have to be repeated again. Many previous machine-learning-based localization models assume that (1) in an offline phase, a lot of labeled training data are available to learn a localization model; (2) the learned localization model is static over time, across space or across devices. Thus it can be used to accurately locate a mobile device online without any adaptation.

However, these assumptions may not hold in many real-world WILPs for several reasons. First, the data distribution may be a function of time, leaving it difficult to apply a trained model to a new scenario at a different time period. Secondly, the data distribution may be a function of space, making it expensive to collect the training data at all locations in a large scale building. Finally, the data distribution may be a function of client device, making the model trained for one type of device (say Cisco) to be invalid when applied to another device (say Intel). To explain these reasons clearly, we show the observations as follows:

- We first consider changing data distribution over time. In a complex indoor building, the environment is always dynamic in nature, caused by unpredictable movements of people, radio interference and signal propagation. Thus, the distributions of RSS values at training and application periods may be significantly different. For example, as shown in Figure 1, the distributions between WiFi signals received at two different time periods can be very different even at a fixed location and by a same device.
- We then consider changing data distribution across space. Since to collect the labeled training data is a very expensive process, especially when the indoor building is at large scale, it is nice that if we can only collect the labeled data from a subarea of a building and unlabeled data from the remaining area. However, in this case, the distributions between labeled data and unlabeled data may be very different.



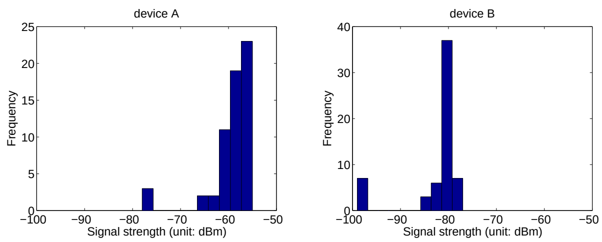
[1] http://en.wikipedia.org/wiki/Activity_recognition




(a) Time Period 1 (b) Time Period 2

Figure 1: Variations of signal-strength histograms in two time periods at the same location from one access point

- Finally, we consider changing data distribution across devices. For example, as shown in Figure 2, the distributions between WiFi signals collected by two different devices can be quite different even at a fixed location and at a same time.




(a) Device A (b) Device B

Figure 2: Variations of signal-strength histograms over two different devices at the same location from one access point.

The above observations may directly cause many traditional machine-learning-based localization models to fail if we use the data collected at one time period, in one subarea or by one device for training and the data collected at another time period, in another subarea or by another device for testing. Therefore, how to reduce the calibration effort of adapting localization models over time, across space or across devices are three major problems of WILP. To solve these problems, we have conducted a series of related transfer learning research at the Hong Kong University of Science and Technology. In particular, in this paper we summarize our work in transferring the localization knowledge across time, space and device. We give an overview of our approaches in our recent works (Pan *et al.* 2007; b; Pan, Kwok, & Yang 2008; Zheng *et al.* 2008b; 2008a) and the provide some experimental results.

## Related Work

### WiFi Localization in Indoor Environments

Received-signal-strength (RSS) based indoor localization and tracking methods have been increasingly popular for WiFi networks (Bahl, Balachandran, & Padmanabhan 2000; Letchner, Fox, & LaMarca 2005). The problem of RSS based indoor localization is to predict locations of a mobile device based on RSS values. Consider a two-dimensional indoor localization problem [2]. A location is represented by $\boldsymbol{\ell} = (x, y)$, where $x$ and $y$ correspond to the x-coordinate and y-coordinate value, respectively. Assume that there are $m$ transmitters, such as Access Points (APs), in an indoor environment, which periodically send out wireless signals. A mobile device can receive signals sent by each of the $m$ APs. Thus, the signals received by a mobile device at a certain location can be represented by a vector $\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_m)^T \in \mathbb{R}^m$. Now the goal of a localization system can be stated formally as estimating the location $\boldsymbol{\ell}_i$ of a mobile device based the signal strength vector $\mathbf{s}_i = (\mathbf{s}_{i1}, \mathbf{s}_{i2}, ..., \mathbf{s}_{im})^T$ received by the mobile device. In general, many localization systems operate in two phases: an *offline* or *training phase* and an *online localization phase*. During the offline phase, a radio mapping function is built from a large amount of signal vectors collected at various locations. In the online phase, the radio mapping function is used to locate a mobile device using its real-time signal vectors. Here the locations of APs are not necessarily known. We call the signal strength vectors with known location information as labeled data, and those without location information as unlabeled data.

Existing approaches to RSS based localization fall into two main categories: radio-propagation-model-based models and machine-learning-based models. The methods in the first category rely on indoor propagation models and the performance is limited because they cannot handle uncertainty (Bahl, Balachandran, & Padmanabhan 2000). Over the last few years, various machine learning approaches have been applied to the indoor localization problem. (Ferris, Hähnel, & Fox 2006; Nguyen, Jordan, & Sinopoli 2005; Pan *et al.* 2005). Two major drawbacks of these approaches are as follows: (1) many of them require a large amount of labeled data for training the models. (2) many of them have a common assumption that the static model learned from training data can be used to accurately estimate locations of mobile device online.

Recently, a few approaches have been proposed for reducing the calibration effort of learning localization models offline, but none have addressed the problem of adapting the models learned in one spatial area to fit another spatial area across an environment. (Ferris, Fox, & Lawrence 2007) applied Gaussian-Process-Latent-Variable models (GP-LVMs) to construct RSS map under an unsupervised learning framework. In this model, an appropriate motion dynamics model needs to be given. (Pan *et al.* a) proposed to apply manifold regularization (Belkin, Niyogi, & Sindhwani 2006) to mobile device tracking in a wireless sensor network under a semi-supervised learning framework. In this model, the labeled training data still need to be uniformly collected over the whole building.

[2] Localization in a three-dimensional environment can be seen as an extension.

Some previous works have also proposed several strategies to adapt localization models over time. The LEASE system (Krishnan *et al.* 2004) utilizes different hardware equipments to solve this problem. LEASE employs a number of stationary emitters and sniffers to obtain up-to-date RSS values for updating the maps. The localization accuracy can only be guaranteed when these additional hardware equipments are deployed in high density. (Bahl, Balachandran, & Padmanabhan 2000) proposed to collect training data in multiple time periods to build models, and then used a probabilistic approach to decide which model should be used online according to the real-time RSS values. However, the calibration effort of such solution is still high.

## Transfer Learning

*Transfer learning* aims to solve the problem when the training data from a source domain and the test data from a target domain follow different distributions or are represented in different feature spaces (Caruana 1997). In the past, transfer learning has been studied in two major contexts. The first context can be referred to as instance-based approach (Dai *et al.* 2007b; Huang *et al.* 2007; Sugiyama *et al.* 2008), where different weights are learned to rank instances in a source domain for using as training data for the target domain. A second context can be referred to as feature-based approach (Ando & Zhang 2005; Argyriou, Evgeniou, & Pontil 2007; Blitzer, McDonald, & Pereira 2006; Raina *et al.* 2007; Dai *et al.* 2007a), which tries to find a common feature set among the different domains that can bridge the two for knowledge transfer. These works include *multi-task learning* (Ando & Zhang 2005; Argyriou, Evgeniou, & Pontil 2007), *multi-domain learning* (Blitzer, McDonald, & Pereira 2006; Dai *et al.* 2007a) and *self-taught learning* (Raina *et al.* 2007).

Transfer learning techniques have been applied successfully in many real world applications, such as learning in real-time strategy game (Sharma *et al.* 2007), text mining (Raina, Ng, & Koller 2006; Dai *et al.* 2007a), natural language processing (Blitzer, McDonald, & Pereira 2006) and so on. However, in the area of machine learning based localization, few work on transfer learning has been done before.

# Transfer Learning for WILP

## Transferring the Localization Models Over Time

Consider transfer learning over time for a two-dimensional WiFi indoor localization problem. We denote the WiFi signal data collected at a time period **0** as $D^0$ and denote WiFi signal data collected at another time period **t** as $D^t$. We assume $D^0$ to be fully labeled whereas $D^t$ to have only a few labeled examples and some unlabeled ones that can be easily obtained by randomly walking through the environment. We collected a lot of labeled data at the time period **0** while only collected a few labeled data at the time period **t**. Our goal is to construct a accurate localization model for the time period **t** from $D^0$ and $D^t$. We propose two solutions to solve this problem according to different situations. If the user trace information is available [3], then a hidden Markov model based solution is preferred. Otherwise, we adopt a manifold co-regularization based solution. Below, we give an overview of our solutions to these problems, where the detailed descriptions can be found in (Pan *et al.* 2007; Zheng *et al.* 2008b).

**Transferring HMM Models over Time** If the *trace* information is available offline and online, then we take it into account to develop a Transferred Hidden Markov Model (TrHMM) to solve this problem. Figure 3 illustrates our idea. We model the prediction problem as a classification problem of discrete location grids. At time 0, we collect RSS data with location labels over the whole area. This step is time consuming, but is done only once. This dataset consists of both the RSS samples at each location and some unlabelled user traces collected as the user walks around the environment in arbitrary trajectories. Based on this data, we train an HMM $\theta_0 = (\lambda_0, A_0, \pi_0)$ for localization at time 0. Basically, $\lambda_0$ is the radio map that connects the RSS values to the locations, $A_0$ is the transition matrix that reflects the way the user moves, and $\pi_0$ is the prior knowledge on the likelihood of where the user is. Note that $\lambda_0$ is changing over time because signal strength varies. $A_0$ can also be changing over time because at different time periods, people may have different activities. For example, at noon, people are more likely to go to canteen for lunch; and at working hours, people are more likely to move within the office area. Therefore, both $\lambda_0$ and $A_0$ need to be adapted to $\lambda_t$ and $A_t$ for a new time period $t$. $\pi_0$ is kept unchanged over time, since in reality, the basic human behavior does not change dramatically. For example, a professor stays at his office each day longer than his walking in a corridor. We propose a solution involved with regression analysis and Expectation Maximization (EM) techniques to adapt the parameters $(\lambda_0, A_0, \pi_0)$ in the time period **0** to parameters $(\lambda_t^{new}, A_t, \pi_t)$ in the time period **t**. The details of the algorithm can be found in (Zheng *et al.* 2008b).

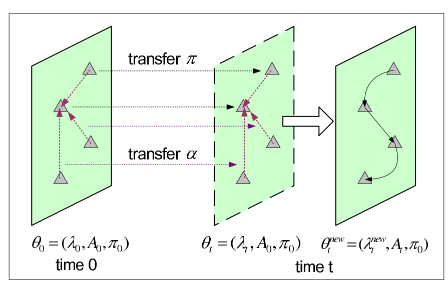



Figure 3: Adapting a localization model from time **0** to time **t** using TrHMM. The triangles denote reference point locations in the area.

[3]The time stamp of each RSS-value record is available.

**Manifold Co-Regularization based Solution** If the *trace* information is unavailable offline or online, then we need to develop another technique to solve this proplem. Observe that although the data distributions are different in high dimensional signal space over time, the data should have a common underlying low-dimensional manifold structure, which can be interpreted as the physical location space. This geometric property is the intrinsic knowledge of the localization problem in a WiFi environment over time. From Figure 4, we can see that although the data in high dimensional signal space have different distributions, they correspond to the same low-dimensional space, the physical space. Thus, if we have some pairwise constraints between the data in different time periods, we can align the high dimensional data in different time periods into a common low-dimensional space, and then learn a localization model that takes both $D^0$ and $D^t$ into account. As ”A”, ”B” and ”C” shown in Figure 4, we can simply place a few reference points, which are additional sensors for recording real-time RSS values, to construct such pairwise, $\{S_A, S'_A\}$, $\{S_B, S'_B\}$ and $\{S_C, S'_C\}$. In (Pan *et al.* 2007), we develop a new manifold co-regularization technique to solve this problem. The details of the algorithm can be found in the paper.

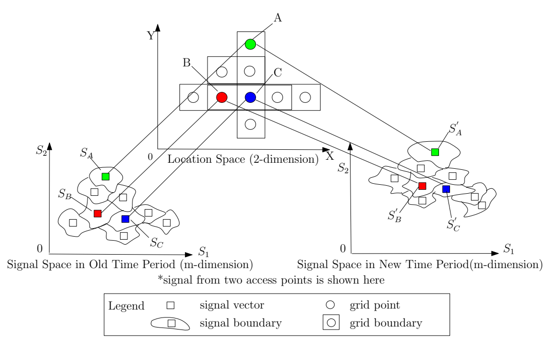



Figure 4: Correlation between location space and two different signal space

## Transferring Localization Models Across Space

We first consider transfer learning across space for a two-dimensional WiFi indoor localization problem. We denote the WiFi signal data collected in an area **A** as $S^a$ and denote WiFi signal data collected in an area **B** as $S^b$. We assume $S^a$ to be fully labeled whereas $S^b$ to have only a few labeled examples and some unlabeled ones that can be easily obtained by quickly walking through the area. The environment as shown in Figure 5 is one of our test beds, whose size is about $35 \times 120\, m^2$. In this test bed, we have collected labeled data in an area we call Area **A** while we only collected less than 10 labeled data in another area we call Area **B**. Our key observation is that there must be some latent knowledge or common structure between $S^a$ and $S^b$, which can be used for propagating the label information across space, when the area **A** and the area **B** being in the same indoor WiFi environment. Our goal is to automatically discover some *shared structure or knowledge* (which can be referred to as $KB$) in the indoor localization domain. We can then incorporate $KB$ and the data $S^a$ and $S^b$ to construct an accurate localization model for the whole space, including both areas **A** and **B**.

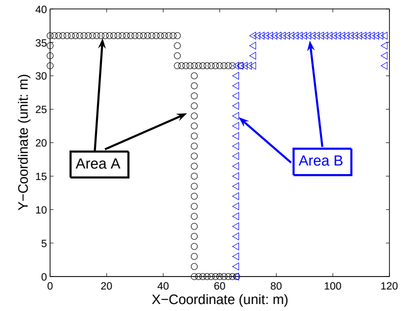



Figure 5: Area **A** and Area **B** in an Environment.

Specifically, our solution for exploiting the data collected in a area for building a localization model for the whole environment through transfer learning consists of two subtasks: (1) Automatically extract the domain knowledge of an indoor environment from the labeled data collected in a area; (2) Incorporate the extracted domain knowledge into a model to propagate the label information to unlabeled data collected in the rest of the environment, and further construct the localization model. For the first subtask, we formulate a quadratically constrained quadratic program (QCQP) optimization problem to discover an underlying semantic manifold of the WiFi signal data. This semantic manifold acts as a bridge that propagates the common knowledge across different areas. For the second subtask, we exploit the common knowledge learnt in the first subtask by taking them as some constraints, and incorporate them to another QCQP optimization problem to estimate labels of the unlabeled data collected in the rest of the environment and then build a mapping function for the whole environment. The details of the algorithm can be found in (Pan *et al.* b).

## Transferring Localization Models Across Devices

Consider transfer learning across devices for a two-dimensional WiFi indoor localization problem. We denote the WiFi signal data collected by a device **A** as $D^a$ and denote WiFi signal data collected by another device **B** as $D^b$. We assume $D^a$ to be fully labeled whereas $D^b$ to have only a few labeled examples and some unlabeled ones that can be easily obtained by quickly walking through the environment. We collected a lot of labeled data by the device **A** while only collected a few labeled data by the device **B**. Observe that although the devices may be different from each other, the learning tasks on these devices are related since they all try to learn a mapping function from a signal space to a *same* location space. This motivates us to model the multi-device localization problem as a multi-task learning problem (Caruana 1997).

Many existing multi-task learning methods assume that the hypotheses learned from the original feature space for related tasks can be similar. This potentially requires the data distributions for related tasks to be similar in the high-dimensional feature space. However, in our multi-device localization problem, the data distributions are expected to be quite different from each other. Therefore, we extend the multi-task learning for multi-device localization problem by only requiring that the hypotheses learned from a *latent* feature space are similar. In other words, we look for appropriate feature mappings, by which we can map different devices' data to a well-defined low-dimensional feature space. In this latent space, new device can benefit from integrating the data collected before by other devices to train a localization model. The details of the algorithm can be found in (Zheng *et al.* 2008a).

## Experiments and Discussion

### ICDM 2007 Data Mining Contest Dataset

Despite intense research in the area of indoor location estimation and activity recognition, there has been a lack of benchmark data with which researchers and practitioners can compare their solutions. We co-organized the first IEEE ICDM Data Mining Contest (IEEE ICDM DMC'07) and published the first such realistic public benchmark data for indoor location estimation from radio signal strength received by a client device from various WiFi APs. We collected the data sets in an academic building in the Hongkong University of Science and Technology, consisting of an area of $145.5\text{m} \times 37.5\text{m}$. Locations were dispersed into $247$ grids, each of which has a size of about $1.5\text{m} \times 1.5\text{m}$. The main task is discrete classification. Regression versions of the tasks were also essential to location estimation, and have been released on the ICDM DMC'07 website [4]. The benchmark data is available (Yang, Pan, & Zheng 2008). The second task of ICDM DMC'07 focused on *Transferring the Learned Knowledge for Indoor Location Estimation*. In this task, the training data were collected at a *different time period* from the test data. To help with the prediction in the case where the training and test data may be from different distributions, *some* labeled test data which can be used as benchmarks are provided. The goal is to predict locations of the test data. The data description details can be found on the ICDM DMC'07 website.

### Experimental Results

We conducted experiments on the ICDM DMC'07 dataset to evaluate our proposed solutions for transferring localization models over time. In addition, we collected new training and test data from different space or by different devices in the same building at the Hong Kong University of Science and Technology to evaluate our proposed solutions for transferring localization models across space and across devices, respectively. The new datasets will be published later as benchmark data as well. Table 1 compares our transfer-learning-based solutions with traditional machine-learning-based solutions over time and over space. The numbers shown in the table are culmulative probabilities at 3-meter error distance [5]. We can see that our proposed transfer-learning-based solutions outperform the localization models that are static . The task that transferring localizations models over devices is even more difficult due to the APs detected by different devices at the same location may be very different. However, our proposed latent multi-task learning based solution can still reduce the average error predicted distance of a localization model from 10 meters to around 5 meters with only a few relabeled data. The detailed experimental results have been analyzed in (Pan *et al.* 2007; b; Zheng *et al.* 2008b; 2008a).

| | Over Time (with trace info.) | Over Time | Over Space |
|---|---|---|---|
| NoTransF | 0.73 | 0.7 | 0.65 |
| TransF | 0.85 | 0.8 | 0.7 |

Table 1: Comparison of culmulative probabilities (error distance = 3m) between our transfer-learning-based solutions and traditional machine-learning-based solutions.

## Conclusion and Future Work

In this paper, we study the problem of exploiting transfer learning algorithms for WILP. We have shown that our proposed transfer-learning-based solutions can transfer localization models over time, across space and across devices, effectively. In the future, we plan to exploit transfer learning into other pervasive computing applications, such as activity recognition from low-level sensory data (Hu & Yang 2008). In real-world scenarios, acquiring the training data for activity recognition of a particular user in a particular environment, with a given set of actions is extremely costly. Therefore, a natural problem that arises is how to transfer the useful knowledge of an activity recognition from a person to to another person or from an environment to another environment. In addition, we also plan to develop a more general transfer learning method for various realistic problems. In (Pan, Kwok, & Yang 2008), we have proposed a novel transfer learning method via dimensionality reduction. The basic idea behind our method is that we try to extract implicit features by which the distributions of the data in different domains are close to each other in the subspace spanned. Based on the extracted features, we can migrate the labeled data across domains. We have applied this method to text

[4] http://www.cse.ust.hk/∼qyang/ICDMDMC07/

[5] Culmulative probabilities at 3-meter is equivalent to the accuracy of predictions (ranges from 0 to 1), where the predictions within 3 meters of the ground truth are all counted as correct predictions, otherwise are counted as incorrect predictions.

classification and WiFi localization problems and achieve encouraging results. We will go on our research and make this method more effective and efficient.

## Acknowledgement

We thank the support of a research grant from NEC-China under project #: NECLC05/06.EG01.

## References

Ando, R. K., and Zhang, T. 2005. A framework for learning predictive structures from multiple tasks and unlabeled data. *Journal of Machine Learning Research* 6:1817–1853.

Argyriou, A.; Evgeniou, T.; and Pontil, M. 2007. Multi-task feature learning. In *Proceedings of the 19th Annual Conference on Neural Information Processing Systems*. 41–48.

Bahl, P.; Balachandran, A.; and Padmanabhan, V. 2000. Enhancements to the RADAR user location and tracking system. Technical report, Microsoft Research.

Belkin, M.; Niyogi, P.; and Sindhwani, V. 2006. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *Journal of Machine Learning Research* 7:2399–2434.

Blitzer, J.; McDonald, R.; and Pereira, F. 2006. Domain adaptation with structural correspondence learning. In *Proceedings of the Conference on Empirical Methods in Natural Language*, 120–128.

Caruana, R. 1997. Multitask learning. *Machine Learning* 28(1):41–75.

Dai, W.; Xue, G.; Yang, Q.; and Yu, Y. 2007a. Co-clustering based classification for out-of-domain documents. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Dai, W.; Yang, Q.; Xue, G.; and Yu, Y. 2007b. Boosting for transfer learning. In *Proceedings of the 24th International Conference on Machine Learning*, 193–200.

Ferris, B.; Fox, D.; and Lawrence, N. 2007. WiFi-SLAM using gaussian process latent variable models. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, 2480–2485.

Ferris, B.; Hähnel, D.; and Fox, D. 2006. Gaussian processes for signal strength-based location estimation. In *Robotics: Science and Systems II*.

Hu, D. H., and Yang, Q. 2008. CIGAR: Concurrent and interleaving goal and activity recognition. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence*.

Huang, J.; Smola, A.; Gretton, A.; Borgwardt, K. M.; and Schölkopf, B. 2007. Correcting sample selection bias by unlabeled data. In *Proceedings of the 19th Annual Conference on Neural Information Processing Systems*.

Krishnan, P.; Krishnakumar, A. S.; Ju, W.-H.; Mallows, C.; and Ganu, S. 2004. A system for LEASE: Location estimation assisted by stationary emitters for indoor RF wireless networks. In *Proceedings of IEEE INFOCOM*.

Letchner, J.; Fox, D.; and LaMarca, A. 2005. Large-scale localization from wireless signal strength. In *Proceedings of the 20th National Conference on Artificial Intelligence*, 15–20.

Nguyen, X.; Jordan, M. I.; and Sinopoli, B. 2005. A kernel-based learning approach to ad hoc sensor network localization. *ACM Transactions on Sensor Networks* 1(1):134–152.

Pan, J. J.; Yang, Q.; Chang, H.; and Yeung, D.-Y. A manifold regularization approach to calibration reduction for sensor-network based tracking.

Pan, S. J.; Shen, D.; Yang, Q.; and Kwok, J. T. Transferring localization models across space. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence*.

Pan, J. J.; Kwok, J.; Yang, Q.; and Chen, Y. 2005. Accurate and low-cost location estimation using kernels. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, 1366–1371.

Pan, S. J.; Kwok, J. T.; Yang, Q.; and Pan, J. J. 2007. Adaptive localization in a dynamic WiFi environment through multi-view learning. In *Proceedings of the 22nd AAAI Conference on Artificial Intelligence*, 1108–1113.

Pan, S. J.; Kwok, J. T.; and Yang, Q. 2008. Transfer learning via dimensionality reduction. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence*.

Raina, R.; Battle, A.; Lee, H.; Packer, B.; and Ng, A. Y. 2007. Self-taught learning: Transfer learning from unlabeled data. In *Proceedings of the 24th International Conference on Machine Learning*, 759–766.

Raina, R.; Ng, A. Y.; and Koller, D. 2006. Constructing informative priors using transfer learning. In *Proceedings of the 23rd International Conference on Machine Learning*, 713–720.

Sharma, M.; Holmes, M.; Santamara, J. C.; Irani, A.; Jr., C. L. I.; and Ram, A. 2007. Transfer learning in real-time strategy games using hybrid CBR/RL. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, 1041–1046.

Sugiyama, M.; Nakajima, S.; Kashima, H.; Buenau, P. V.; and Kawanabe, M. 2008. Direct importance estimation with model selection and its application to covariate shift adaptation. In *Proceedings of the 20th Annual Conference on Neural Information Processing Systems*.

Yang, Q.; Pan, S. J.; and Zheng, V. W. 2008. Estimating location using Wi-Fi. 23(1):8–13.

Yin, J.; Yang, Q.; and Ni, L. M. 2005. Adaptive temporal radio maps for indoor location estimation. In *Proceedings of the 3rd IEEE International Conference on Pervasive Computing and Communications*.

Zheng, V. W.; Pan, S. J.; Yang, Q.; and Pan, J. J. 2008a. Transferring multi-device localization models using latent multi-task learning. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence*.

Zheng, V. W.; Yang, Q.; Xiang, W.; and Shen, D. 2008b. Transferring localization models over time. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence*.