Total Variation Regularization for Training of Indoor Location Fingerprints

Duc A. Tran
University of Massachusetts Boston
100 Morrissey Blvd
Boston, MA
duc.tran@umb.edu

Phong Truong
University of Massachusetts Boston
100 Morrissey Blvd
Boston, MA
phong.truong001@umb.edu

ABSTRACT

Location fingerprinting is a common approach to indoor localization. For good accuracy, the training set of sample fingerprints, each mapping a fingerprint to a location, should be sufficiently large to be well-representative of the environment in terms of both spatial coverage and temporal coverage. Unfortunately, the task of collecting these samples can be tedious and labor-intensive because one must label each location that is being surveyed. On the other hand, fingerprints without location information are abundant and can easily be collected and so recent studies have tried to capitalize on these unlabeled fingerprints to improve the training set. The paper investigates how this goal can be achieved via graph regularization based on Total Variation (TV). TV is highly effective for semi-supervised learning in image processing but it is not clear whether its success can be transferred to indoor location fingerprinting.

Categories and Subject Descriptors

H.1.0 [**Information Systems**]: Models and Principles

Keywords

Location fingerprint, localization, total variation

1. INTRODUCTION

Location information is valuable to a myriad of indoor applications of wireless networks. GPS is the most popular way to get location information but does not work indoors. Sans GPS, an alternative approach is to leverage the correlation between received signal strength (RSS) and distance[21]. If we have a number of reference points (RPs), e.g., Wi-Fi access points [1], FM broadcasting towers [6], or cellular towers [20], we can locate a device by estimating the distances between this device to the RPs based on RSS ranging and then using multi-lateration to compute the location. RSS ranging, however, is highly inaccurate due to noise interference [21].

ACM MiSeNet'13 October 4, 2013, Miami, Florida, USA
Copyright 2013 ACM 978-1-4503-2367-3/13/10 ...$15.00.
http://–enter the whole DOI string from rightsreview form confirmation.

Arises the fingerprint approach, which is range-free. This approach consists of two phases: training (offline) and positioning (online). In the offline phase, a number of sample locations are surveyed to build a map of "fingerprints", each corresponding a surveyed location to a vector of RSS values received at this location from a set of RPs. In the online phase, when we need to compute a location in real time, the RSS vector of the device is evaluated against the fingerprint map to find the best location match. Most often, as employed in [1, 13, 23], the centroid of the locations corresponding to one or more nearest fingerprints in the fingerprint space is used as the estimate for the device's location. While this kNN-based method compares RSS values directly, one can use a model-based learning method, e.g., [17, 4], to compute a functional dependence between a fingerprint and a location.

Despite its simplicity, the fingerprint approach is limited by the quality of the training data. The training data should be sufficiently large to be well-representative of the environment, both spatially and temporally. For a large area, many locations need to be surveyed to ensure good spatial coverage and many fingerprint readings need to be measured at each sample location to ensure good temporal coverage (due to the dynamics of the environment over time).

As the calibration task is tedious and labor-extensive, it is often the main bottleneck to localization accuracy. Consequently, various efforts [15, 14, 16, 10] aim to reduce the calibration cost by a semi-supervised learning approach [5], utilizing unlabeled fingerprints to supplement a given (small) training set of labeled fingerprints. A labeled fingerprint is one that is measured at a known location. Unlabeled fingerprints are abundant because they can easily be obtained for a mobile device without manual location labeling.

Our research is in the same sprit with the aforementioned works, with the exception that we use Total Variation (TV) regularization for semi-supervised learning of the unlabeled fingerprints. TV permits sharper edges near the decision boundaries whereas other regularization methods such as Laplacian used in manifold regularization [2] penalize too much gradients on edges. Therefore, TV is highly successful to image de-noising and restoration tasks as it does not smoothen edges unnecessarily in the picture. Our rationale for exploring TV in indoor localization is because (1) it is observed that the fingerprint space exhibits a degree of smoothness (nearby locations should correspond to similar fingerprints), hence the rationale for using a regularization method, and (2) whether the success of TV in image processing may apply to indoor localization is not clear as they are in different signal spaces; this paper is an effort to investigate this question We explore the applicability of TV to improving the training quality of Wi-Fi fingerprint datasets obtained from two different indoor testbeds.

The rest of the paper is organized as follows. §2 provides a brief survey of the related work. §3 describes how we apply Total Variation to our semi-supervised learning problem. Evaluation results are discussed in §4. The paper is concluded in §5 with pointers to our future work.

2. RELATED WORK

Radar [1] is the world's first Wi-Fi RSS-based indoor positioning system, which demonstrates the viability of using RSS information to locate a wireless device. This system works using a radio map, a lookup table that maps building locations to the corresponding RSS fingerprints empirically observed at these locations. The reference points are the Wi-Fi access points within the user's Wi-Fi range. The radio map is searched to find the closest RSS reading and the corresponding location will be used as the estimate for the user's location. MoteTrack [13] is a location fingerprint technique largely similar to Radar, but aimed at improving the robustness of the system through a decentralized approach to estimating the location. Instead of location computation's being performed at a central back-end server, this task is distributed among the reference nodes. Given a fingerprint at an unknown location, the location can be computed as the centroid of a number of nearest fingerprints in the radio map, using the Manhattan distance for similarity measure. More weight is given to a training fingerprint that is nearer.

Radar and MoteTrack represent the fingerprint approach where kNN is used to determine the location. One can also employ a model-based learning approach to relate a fingerprint to a location. This approach is motivated by the stochastic nature of fingerprint information: Fingerprints at the same location may vary over time and, due to presence of physical obstacles, they may have identical values even at two distant locations. Consequently, it is argued that location estimation may be more accurate if an intrinsic functional dependence between a fingerprint and a location exists and is utilized. This dependence can be represented probabilistically using Bayesian inference [17, 24] or non-probabilistically using an Artificial Neural Network (ANN) [11] or a Support Vector Machine (SVM) [4, 22, 9].

To improve training quality especially when there are only a small number of sample fingerprints, a viable approach is to harness unlabeled fingerprints as a supplement to the original ones, by using a semi-supervised learning tool to propagate the labels for the unlabeled fingerprints based on their similarity with the labeled. This is possible because of the smoothness in the fingerprint space, as empirically demonstrated in [15]. Pan et al. [15, 14] apply manifold regularization originally proposed in [2], where the semi-supervised learning problem is formulated as a minimization problem on a regularized risk functional with a regularization term reflecting the intrinsic manifold structure of the fingerprints. Pulkkinen et al. [16] employs the Isomap algorithm [19] to project the fingerprints onto a low-dimension manifold representation with minimal Euclidean distance distortion, and then, based on the manifold topology and the locations of labeled fingerprints, propose a calibration method to map a manifold coordinate to a geographical coordinate. In both of these works, the learning is applied on a k-nearest-neighbor or ε-ball neighborhood graph of fingerprints (labeled and unlabeled). The recent work by Zhu et al. [10] introduces a regularization term based on the l^1-graph [7].

Our research in this paper also applies a graph regularization framework for learning the location labels for the unlabeled fingerprints, but our regularization term is based on the Total Variation (TV) norm. TV is on of the most successful tool for image restoration tasks such as denoising and inpainting [18] and recently been extended to the supervised learning settings on high dimensional data [12].

3. TV FRAMEWORK

Suppose that we have obtained a collection of $n = l+u$ fingerprints $\{x_1, x_2, ..., x_l, x_{l+1}, ..., x_{l+u}\}$, l of which are labeled, $\{x_1, x_2, ..., x_l\}$, and u unlabeled, $\{x_{l+1}, ..., x_{l+u}\}$. Each fingerprint is a m-dimensional point, $x_i \in \mathbb{R}^m$, where m is the dimensionality of the fingerprint space (e.g,. the number of Wi-Fi access points whose signal strengths are used for fingerprinting). We denote by $y_i \in \mathbb{R}^d$ the location corresponding to x_i, where d is the dimensionality of the location space. The labeled fingerprints are those in the original training set (fingerprint map) obtained during the location survey; hence, $\{y_1, y_2, ..., y_l\}$ are known. We need to predict the values for $\{y_{l+1}, y_{l+2}, ..., y_{l+u}\}$ for the unlabeled fingerprints. Once this task is completed, the extended fingerprint map of size $l + u$ will be used instead of the original map of size l. For example, if kNN is used during the positioning phase, the extended map will be evaluated to find the nearest locations corresponding to the test fingerprint.

Without loss of generality, we assume that the coordinate range for x_i and y_i is [0, 1]; that is, $x_i \in \Omega = [0, 1]^m$, $y_i \in [0, 1]^d$. Also, for ease of presentation, we assume for now that $d = 1$ and so y_i is a real number. We will extend our framework to the case $d > 1$ later.

3.1 Manifold Regularization

Supervised learning can be formulated as a Tikhonov regularization problem. We need to estimate an unknown real-valued function $f : \Omega \to \mathbb{R}$ that relates a point x_i (fingerprint in our case) to a value $f_i = f(x_i)$ (location in our case) such that

$$\min_f \quad \sum_{i=1}^{n} L(f_i, y_i) + \lambda \|f\|_K^2,$$

where L denotes a loss function, e.g., hinge loss $\max(0, 1 - y_i f_i)$ used in Support Vector Machines (SVM) or squared loss $(f_i - y_i)^2$ in Regularized Least Squares (RLS). The second term represents the smoothness penalty with respect to a standard kernel K (note that f must belong to a RKHS family of functions associated with kernel K).

Extending this framework to semi-supervised learning, Belkin et al. [2] propose the Manifold Regularization method, which includes an extra Laplacian smoothing term reflecting the smoothness with respect to an intrinsic manifold, as follows

$$\min_f \quad \frac{1}{l} \sum_{i=1}^{l} (f_i - y_i)^2 + \lambda \|f\|_K^2 + \gamma \underbrace{\sum_{i,j=1}^{n} w_{ij} |f_i - f_j|^2}_{f^T L f}.$$

Here, a nearest-neighbor undirected self-loop-less graph W needs be constructed first, where each vertex represents a sample and weight w_{ij} the similarity between samples i and

j (non-zero if i and j are connected by an edge and zero otherwise). The Laplacian, L, of this graph provides a natural intrinsic measure of smoothness (the Dirichlet's energy $f^T Lf$ shown in the last term).

3.2 TV Regularization

Our goal is to explore how effective TV can be to semi-supervised learning with our fingerprint collection. Manifold regularization has been applied in [15]. Our work, which is the first that explores TV, is inspired by the success of TV in image processing tasks [12] and recent evidences showing its promise in semi-supervised classification [3]. Our problem is to predict the locations directly not based on classification.

In the TV framework for semi-supervised learning, the minimization problem is

$$\min_f \quad \frac{1}{l}\sum_{i=1}^{l}(f_i - y_i)^2 + \lambda\|f\|_K^2 + \gamma TV_W \tag{1}$$

where TV_W is the total variation of function f on graph W. By definition, the TV of a continuous function $f : \Omega \to \mathbb{R}$ is

$$TV[f] = \int_\Omega \|\nabla f\| \; dx.$$

Here, ∇f is the gradient of function f and dx the area element of the continuous domain Ω. We can extend this concept to weighted graphs as follows. Given a graph W, the total variation of a real-valued scalar function f defined on its vertices is the sum of local total variation at each and every vertex,

$$TV_W = \sum_{i=1}^{n} \|\nabla f(i)\|_{L^p(w)}$$

where the local total variation at vertex i is the weighted L^p-norm of the gradient at this vertex. The gradient of function f at vertex i is

$$\nabla f(i) = \begin{pmatrix} f_1 - f_i \\ f_2 - f_i \\ \dots \\ f_j - f_i \\ \dots \\ f_n - f_i \end{pmatrix}$$

and so,

$$\|\nabla f(i)\|_{L^p(w)} = \left(\sum_{j=1}^{n} w_{ij}|f_j - f_i|^p\right)^{1/p}.$$

The graph TV above is a generalization of that defined in [3] and [8]. The case $p = 1$ corresponds to the graph TV used by Bresson and Zhang [3],

$$TV_W = \sum_{i,j=1}^{n} w_{ij}|f_j - f_i|, \tag{2}$$

and the case $p = 2$ corresponds to the graph TV used by Elmoataz et al. [8],

$$TV_W = \sum_{i=1}^{n} \sqrt{\sum_{j=1}^{n} w_{ij}(f_j - f_i)^2}. \tag{3}$$

3.3 Graph Construction

The formation of the Manifold Regularization and TV Regularization frameworks requires the construction of a similarity graph W connecting the samples x_i. This graph is usually built as a kNN-graph or a ϵ-ball-graph.

- kNN-graph: A vertex is connected to the k nearest vertices based on a pairwise distance measure (e.g., Euclidean or Manhattan). The weight w_{ij} is non-zero only for adjacent vertices x_i and x_j. A Gaussian RBF of the distance is often used for this weight.

- ϵ-ball-graph: A vertex is connected to those among the other vertices whose Euclidean distance is bounded by threshold ϵ. The weight is assigned to an edge as in kNN-graph.

3.4 TV Training in Our Study

To semi-train our fingerprint collection, in this paper, we build graph W as a (Euclidean) kNN neighborhood graph where value of k is chosen to be the minimum one that results in a connected graph. The edge weight w_{ij} between neighbor fingerprints x_i and x_j is assigned the following value

$$w_{ij} = \exp\left(-|x_i - x_j|^2\right)$$

Furthermore, for the minimization problem (1) we use the TV defined according to Formulae (3) and use the algorithm as proposed in [8] to solve this problem. Our purpose is to learn the locations for the unlabeled fingerprints and as such we do not change the locations already labeled for the labeled fingerprints; hence, the loss term is dismissed. Also, since we limit our interest in this paper to investigating the significance of TV in enriching the training data, we focus only on minimizing this term.

Our algorithm to estimate the locations for the unlabeled fingerprints works iteratively as follows:

1. Initial step: for $i, j \in [1, n]$

$$f_i^{(0)} = \begin{cases} y_i & \text{if } i \le l \\ 1/2 & \text{otherwise} \end{cases} \tag{4}$$

$$\gamma_{ij}^{(0)} = w_{ij} \tag{5}$$

2. Iterative step: for $i, j \in [1, n]$

$$f_i^{(t+1)} = \begin{cases} y_i & \text{if } i \le l \\ \frac{\sum_{j=1}^{n} \gamma_{ij}^{(t)} f_j^{(t)}}{\sum_{j=1}^{n} \gamma_{ij}^{(t)}} & \text{otherwise} \end{cases} \tag{6}$$

$$\gamma_{ij}^{(t+1)} = \frac{w_{ij}}{\|\nabla f^{(t)}(i)\|_{L^2(w)}} + \frac{w_{ij}}{\|\nabla f^{(t)}(j)\|_{L^2(w)}} \tag{7}$$

3. Stop when $|f^{(t+1)} - f^{(t)}|$ is less than some small threshold τ (predetermined). The value of $f_i^{(t)}$ $(i > l)$ will be used as the estimated location for each unlabeled fingerprint x_i.

The above algorithm assumes that y_i is 1D. For 2D or a higher dimension ($d > 1$) of the location space, this algorithm is applied separately for learning each coordinate.

4. EVALUATION

We evaluate with a dataset obtained from an indoor experiment used in [4] (University of Trento), containing a

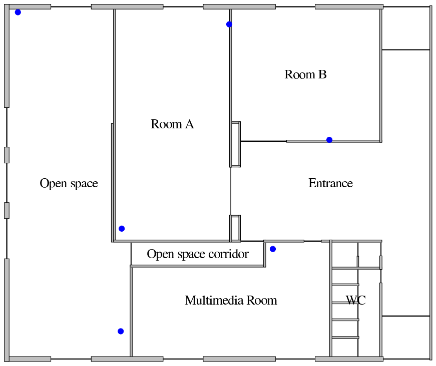


Figure 1: Trento dataset's map: 30m×20m (courtesy of [4])

collection of 257 RSSI fingerprints at 257 sample locations in a WLAN with six Wi-Fi access points (Figure 1). The sample locations are regular-grid points of the floor. Each fingerprint is measured at a sample location by a person carrying a PDA, as a receiver receiving signals from the access points. The PDA always points north. A random half $Train$ of this collection (128 samples) is used for training and the other half $Test$ (129 samples) for testing purposes. Out of the training samples, we randomly create two groups of samples: the labeled group L (with the location labels intact) and the unlabeled group U (with the location labels removed). It is noted that $L, U \subset Train$ and $L \cap U = \emptyset$. The size of L is set to be 10%, 20%, ..., or 90% of $|Train|$ and, given L, the size of U is set to be 10%, 20%, ..., or 100% of $|Train - L|$. For each combination of these sizes, the average location when tested with $Test$ is averaged over 10 times running the simulation (with random generations of U and L). 1-NN is used for testing; that is, given a test fingerprint, its estimated location is the location of the nearest fingerprint in the fingerprint map.

The performance metrics are location error average and max. We compare TV with (1) "Original": the original labeled set L is used as the fingerprint map; (2) "Combine": the set $L \cup U$ where the original label is known for every fingerprint, as the fingerprint map; and (3) "Manifold": the Manifold Regularization method where the Dirichlet's energy is minimized instead of TV.

Figure 2 plots the average error for various cases of $|L|$ and $|U|$. We observe the following patterns:

- Regardless of the size of the labeled set, manifold and TV regularization frameworks tend to be increasingly effective as the size of the unlabeled set increases.
- When the labeled set is small (e.g., 10%), regularization does not help. Only when the labeled set gets sufficiently large (e.g., 60%), we start to see its effectiveness. This finding is understandable because a small labeled set offers too little information to be useful for the training.
- When the labeled set gets sufficiently large for the regularization to be effective, both TV and Manifold can approach the accuracy of Combine as the size of the unlabeled set gets larger (e.g., > 80%). In other words, these regularization methods do have an excellent prediction accuracy close to being perfect.
- Manifold is consistently more accurate than TV. This is different from the observation in the area of image processing where TV is known to be better. This suggests that, unlike images which often have edges, the fingerprint space may not exhibit "edges" of fingerprints (i.e., a path in the fingerprint graph) located at a small cluster of locations isolated from the rest of locations. This could be due to the penetrable-ness of the Wi-Fi signal in the indoor area.

Similar findings are observed for the max error, as seen in Figure 3. Both Manifold and TV can improve the accuracy of the fingerprint map in terms of average and max error.

5. CONCLUSIONS

We have investigated the use of TV regularization as a semi-supervised learning tool for improving the quality of the training set used in indoor location fingerprinting. We have shown that while TV can be effective for this purpose, it is not as good as the widely used Manifold Regularization. However, this finding is meant to be suggestive rather than conclusive as our work in this paper remains preliminary. Our evaluation has been done with only a small dataset and as the next step we will extend our work to investigate with other datasets, different weight functions, and more comprehensive evaluation configurations.

Acknowledgments

Our work was supported in part by the NSF award CNS-1116430. Any opinions, findings and conclusions or recommendations expressed in this material are ours and do not necessarily reflect those of the NSF.

6. REFERENCES

[1] P. Bahl and V. N. Padmanabhan. Radar: An in-building rf-based user location and tracking system. In *INFOCOM*, pages 775–784, 2000.

[2] M. Belkin, P. Niyogi, and V. Sindhwani. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *J. Mach. Learn. Res.*, 7:2399–2434, Dec. 2006.

[3] X. Bresson and R. Zhang. Tv-svm: Total variation support vector machine for semi-supervised data classification. *CoRR*, abs/1210.0699, 2012.

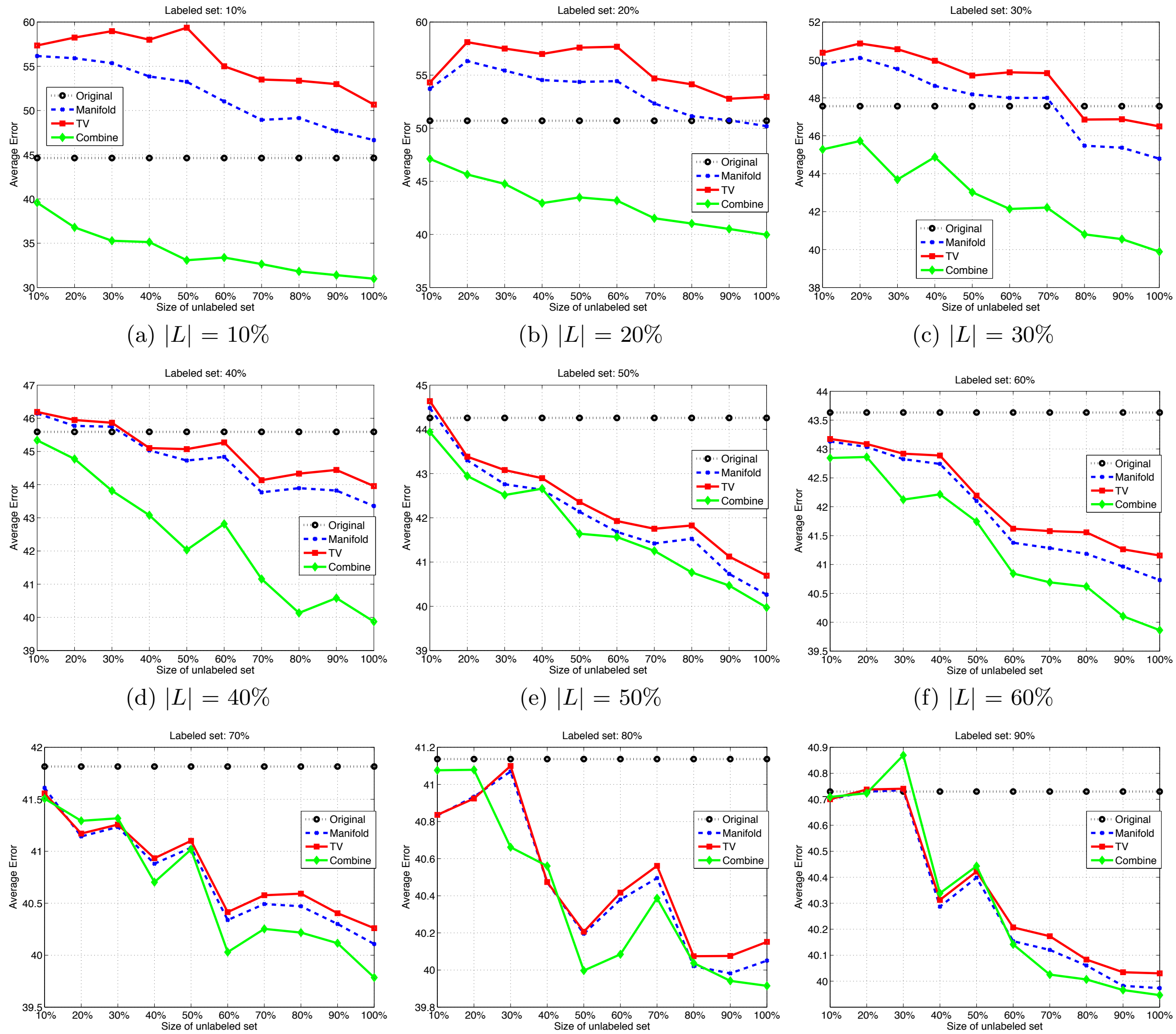
[4] M. Brunato and R. Battiti. Statistical learning theory for location fingerprinting in wireless lans. *Comput. Netw.*, 47(6):825–845, Apr. 2005.

[5] O. Chapelle, B. Schlkopf, and A. Zien. *Semi-Supervised Learning.* The MIT Press, 1st edition, 2010.

[6] Y. Chen, D. Lymberopoulos, J. Liu, and B. Priyantha. Fm-based indoor localization. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, MobiSys '12, pages 169–182, New York, NY, USA, 2012. ACM.

[7] B. Cheng, J. Yang, S. Yan, Y. Fu, and T. S. Huang. Learning with l1-graph for image analysis. *Trans. Img. Proc.*, 19(4):858–866, Apr. 2010.

[8] A. Elmoataz, O. Lezoray, and S. Bougleux. Nonlocal discrete regularization on weighted graphs: A framework for image and manifold processing. *Trans. Img. Proc.*, 17(7):1047–1060, July 2008.



(a) $|L| = 10\%$ (b) $|L| = 20\%$ (c) $|L| = 30\%$

(d) $|L| = 40\%$ (e) $|L| = 50\%$ (f) $|L| = 60\%$

(g) $|L| = 70\%$ (h) $|L| = 80\%$ (i) $|L| = 90\%$

Figure 2: Average Location Error

[9] C. Figuera, J. L. Rojo-Álvarez, M. Wilby, I. Mora-Jiménez, and A. J. Caamaño. Advanced support vector machines for 802.11 indoor location. *Signal Process.*, 92(9):2126–2136, Sept. 2012.

[10] Y. jia Zhu, Z. liang Deng, and H. Ji. Indoor localization via l1-graph regularized semi-supervised manifold learning. *The Journal of China Universities of Posts and Telecommunications*, 19(5):39 – 91, 2012.

[11] C. Laoudias, D. G. Eliades, P. Kemppi, C. G. Panayiotou, and M. M. Polycarpou. Indoor localization using neural networks with location fingerprints. In *Proceedings of the 19th International Conference on Artificial Neural Networks: Part II*, ICANN '09, pages 954–963, Berlin, Heidelberg, 2009. Springer-Verlag.

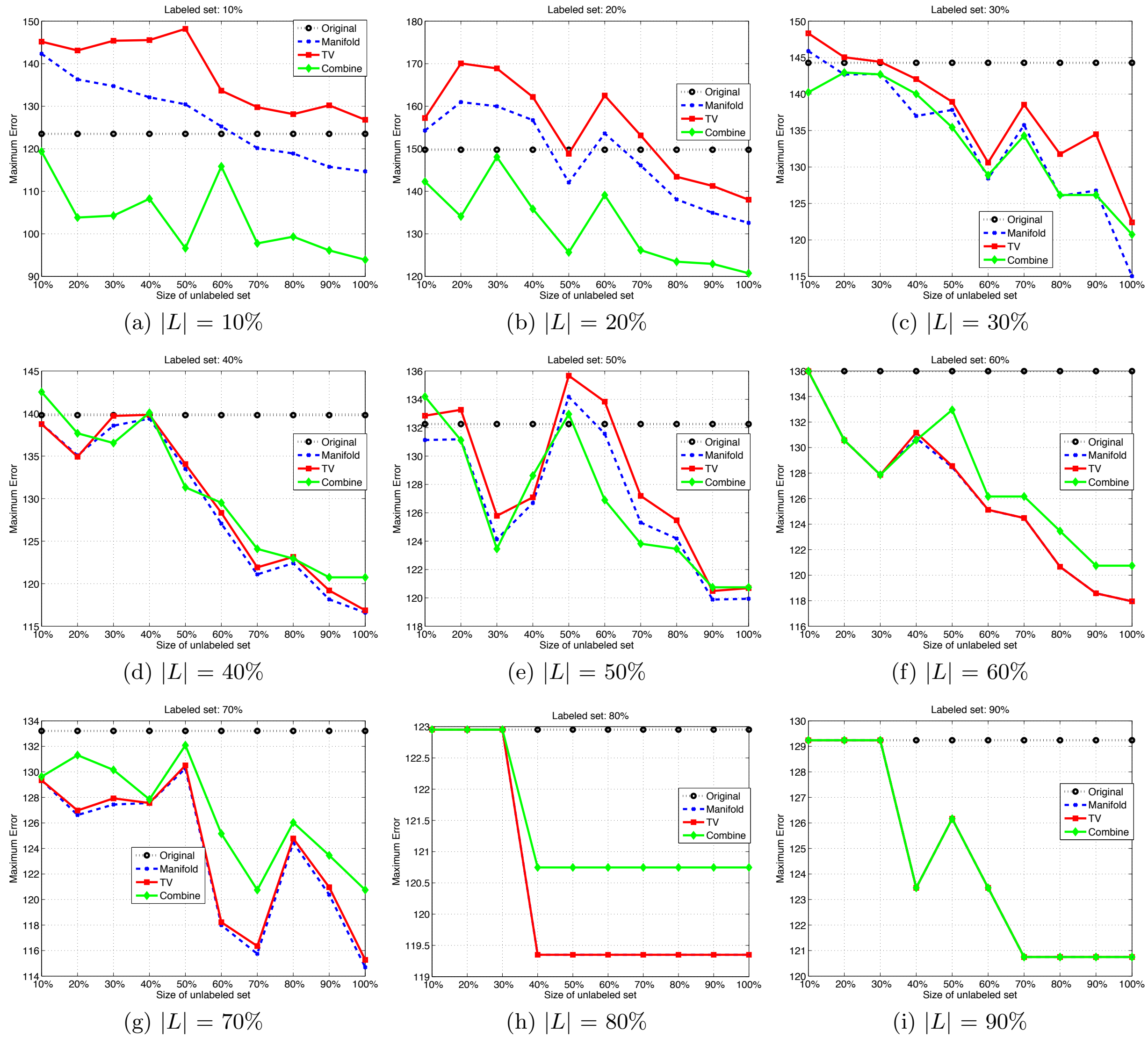
[12] T. Lin, H. Xue, L. Wang, and H. Zha. Total variation and euler's elastica for supervised learning. In *ICML*, 2012.

[13] K. Lorincz and M. Welsh. Motetrack: a robust, decentralized approach to rf-based location tracking. *Personal Ubiquitous Comput.*, 11(6):489–503, Aug. 2007.

[14] J. J. Pan and Q. Yang. Co-localization from labeled and unlabeled data using graph laplacian. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence*, pages 2166–2171, Hyderabad, India, 2007.

[15] J. J. Pan, Q. Yang, H. Chang, and D. Y. Yeung. A manifold regularization approach to calibration reduction for sensor-network based tracking. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence*, pages 988–993, Boston, United States, 2006.

[16] T. Pulkkinen, T. Roos, and P. Myllymäki. Semi-supervised learning for wlan positioning. In *Proceedings of the 21th international conference on*



(a) $|L| = 10\%$ (b) $|L| = 20\%$ (c) $|L| = 30\%$

(d) $|L| = 40\%$ (e) $|L| = 50\%$ (f) $|L| = 60\%$

(g) $|L| = 70\%$ (h) $|L| = 80\%$ (i) $|L| = 90\%$

Figure 3: Max Location Error

Artificial neural networks - Volume Part I, ICANN'11, pages 355–362, Berlin, Heidelberg, 2011. Springer-Verlag.

[17] T. Roos, P. Myllymäki, H. Tirri, P. Misikangas, and J. Sievänen. A probabilistic approach to WLAN user location estimation. *International Journal of Wireless Information Networks*, 9(3):155–164, July 2002.

[18] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, Nov. 1992.

[19] J. B. Tenenbaum, V. D. Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 2000.

[20] A. Varshavsky, E. de Lara, J. Hightower, A. LaMarca, and V. Otsason. Gsm indoor localization. *Pervasive Mob. Comput.*, 3(6):698–720, Dec. 2007.

[21] K. Whitehouse, C. Karlof, and D. Culler. A practical evaluation of radio signal strength for ranging-based localization. *SIGMOBILE Mob. Comput. Commun. Rev.*, 11(1):41–52, Jan. 2007.

[22] C.-L. Wu, L.-C. Fu, and F.-L. Lian. WLAN location determination in e-home via support vector classification. In *Networking, Sensing and Control, 2004 IEEE International Conference on*, volume 2, pages 1026–1031 Vol.2, 2004.

[23] Z. Yang, C. Wu, and Y. Liu. Locating in fingerprint space: wireless indoor localization with little human intervention. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, Mobicom '12, pages 269–280, New York, NY, USA, 2012. ACM.

[24] M. Youssef, A. Agrawala, and A. U. Shankar. Wlan location determination via clustering and probability distributions. In *In IEEE PerCom 2003*, 2003.