

# Multi-Directional Weighted Interpolation for Wi-Fi Localisation

Dale Bowie, Jolon Faichney and Michael Blumenstein  
School of Information and Communication Technology  
Griffith University, Gold Coast, Australia  
Dale.Bowie@griffithuni.edu.au  
{j.faichney, m.blumenstein}@griffith.edu.au

## Abstract

The rise in popularity of unmanned autonomous vehicles (UAV) has created a need for accurate positioning systems. Due to the indoor limitations of the Global Positioning System (GPS), research has focused on other technologies which could be used in this landscape with Wi-Fi localisation emerging as a popular option. When implementing such a system, it is necessary to find an equilibrium between the desired level of final precision, and the time and money spent training the system. We propose Multi-Directional Weighted Interpolation (MDWI), a probabilistic-based weighting mechanism to predict unseen locations. Our results show that MDWI uses half the number of training points whilst increasing accuracy by up to 24%.

## 1 Introduction

Methods of human and robot localisation are fast becoming important areas of research. The Global Positioning System (GPS) has widely been accepted as a means of determining a device's location outdoors. However, there are a number of limitations with GPS that make it undesirable for indoor use: the inability to detect levels, poor location quality because it requires direct contact with satellites, and poor location accuracy because indoor localisation requires a higher level of precision.

As a result, various methods of indoor localisation have been evaluated in literature. The methods differ by the technology as well as their applications; papers have proposed the use of infrared, radio frequency identifier (RFID), Bluetooth, ZigBee and Wi-Fi [2, 6, 10, 12], for various applications of human or robot tracking or navigation [7, 9]. Wi-Fi is seen as a desirable option due to its low setup costs and high availability [1, 7, 9].

Traditionally Wi-Fi localisation works by training a system on signal strength readings obtained from wireless networks inside a building. Readings from a

single location are collated together to form a fingerprint. When deployed, the system will take a device's current fingerprint and compare it to all existing fingerprints to find the best-matching location [1]. Systems using other technologies have followed similar training and positioning phases [12].

As one of the main limitations of indoor localisation schemes is the physical training time required, efforts are being made to attempt to reduce this [3, 8]. Of course, reducing the training space may seem like a simple and effective measure to reduce costs, however they generally result in lower accuracy rates [5]. In this paper we propose Multi-Directional Weighted Interpolation (MDWI) which utilises the probabilities produced by a localisation algorithm to recognise a location outside of the training space.

## 2 Related work

The RADAR system, published by Microsoft Research in 2000 [1], is generally recognised as the founding paper in the field of Wi-Fi localisation, with 5,700 citations to date [4]. The authors gathered Wi-Fi data in one of two ways. In the first approach, samples were manually taken at a number of locations facing different directions, totalling a training space of at least 5,600 readings. The second approach used a signal propagation model to predict signal strengths at those same locations, based on the known locations of access points. The system had consistent results with a median error distance of between 2 and 3 metres, with the first method performing better overall.

In 2002, Youssef et al. [11] proposed the application of probabilistic techniques, such as Bayesian Inference, to the problem of indoor localisation. Based on prior training records as well as a sample reading obtained from a device, they calculated the probability for each possible location. The output was the location with the highest probability, above a pre-defined confidence threshold.

As part of a wider study in 2004, Elnahrawy et al. [3] proposed the Interpolated Map Grid (IMG). IMG utilised a small sample of training fingerprints to interpolate the expected signal strength at various intermediary locations. IMG essentially combines the two methods proposed in the RADAR paper.

In 2006, Li et al. [8] evaluated the use of inverse distance weighting (IDW) and universal kriging (UK). As the name suggests, IDW applies weights inversely proportionate to the distance between the known points. UK generates a polynomial model representative of the known points, and uses this to interpolate intermediary points. The researchers' results reduced the error distance in all cases, with UK consistently outperforming both the standard and IDW methods. When 16 access points were used, UK reduced the original error distance of approximately 2.5 metres down to 1.5 metres.

In their 2007 Signal Strength Difference (SSD) paper Hossain et al. [5] adopted Li et al.'s approach, however, they used a weighted linear regression technique to develop their training set. Equations were created for each pair of access points and intermediary points. Hossain et al.'s test results show the usage of these intermediary points consistently improves the average error

Table 1: Comparison of techniques

Technique	Signal strength	Inverse distance	Probability	Infinite locations
RADAR [1]	No	No	No	No
Prob. Cluster [11]	No	No	Yes	No
IMG [3]	Yes	No	No	No
IDW [8]	No	Yes	No	No
SSD [5]	No	Yes	No	No
MDWI	No	No	Yes	Yes

distance, with a maximum improvement of approximately 3.5 metres.

The techniques mentioned above are compared in table 1 with our technique, MDWI. The authors of prior works propose various methods of interpolation which predict new locations prior to runtime. One of the main disadvantages of performing interpolation prior to runtime is it still enforces a limited state space when deployed. Some propose the use of averaging the  $n$  best-matching locations, however, this disregards the inherent probabilities and potentially results in unlikely locations being included in the final result.

### 3 Multi-Directional Weighted Interpolation

In this paper we propose Multi-Directional Weighted Interpolation which provides greater precision in localisation without an increase in the training time or associated costs. We utilise the probabilities output from an algorithm, such as a Bayesian Network, as weights between the  $n$  best-matching, known fingerprints. In general this will return an intermediary location not included in the original training set.

The basic procedure can be expressed as a series of formulas which were derived from a weighted centre of gravity algorithm. Equations 1 and 2 are equivalent to calculating the dot product of the individual coordinate vectors and the probability vector. The result is the estimated  $x$  and  $y$  coordinates for the current location when given the coordinates of  $n$  best matches (i.e.  $x_i$  and  $y_i$ ) as well as their probabilities ( $p_i$ ).

$$x = \sum_{i=1}^n x_i p_i \tag{1}$$

$$y = \sum_{i=1}^n y_i p_i \tag{2}$$

These formulas can be used for any number of matches. Traditional interpolation would be where  $n = 2$ . Examples follow which show  $n = 3$  and  $n = 4$ .

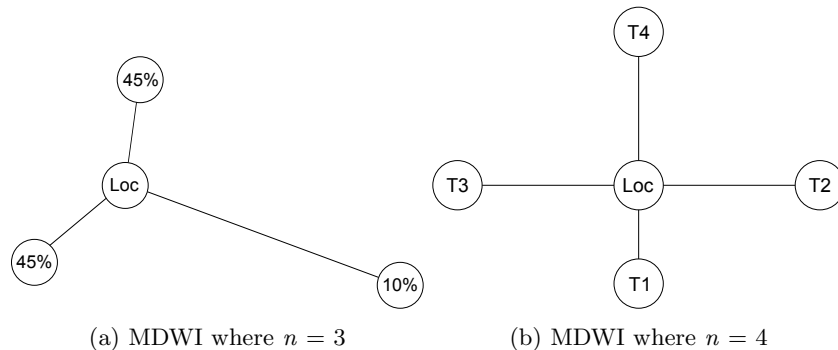


Figure 1: Examples of MDWI

Take the example shown in figure 1a. The three outer nodes are known fingerprints while the central node is the actual location of a user. Some sample probabilities from a Bayesian Network are shown inside the outer nodes. In traditional approaches, either of the two nodes with 45% probability would be chosen. In our approach these probabilities will be used as weights determining a location in the centre of the three fingerprints, much closer to the user’s actual location.

Using figure 1b as an example, we shall now explicitly utilise the formulas. Suppose we have 4 training locations which are labelled T1 through to T4. Their coordinates are (3, 4), (5, 3), (1, 3) and (3, 1), respectively. The coordinates of the actual location are (3, 3). In an ideal scenario, where the algorithm accurately calculates these probabilities, we will have a probability of 40% for T1, and 20% for each of the remaining nodes. From this information we can calculate the predicted location:

$$\begin{aligned}
 x &= \sum_{i=1}^n x_i p_i & y &= \sum_{i=1}^n y_i p_i \\
 &= (3 \times 0.4) + (5 \times 0.2) & &= (4 \times 0.4) + (3 \times 0.2) \\
 &\quad + (1 \times 0.2) + (3 \times 0.2) & &\quad + (3 \times 0.2) + (1 \times 0.2) \\
 &= 1.2 + 1 + 0.2 + 0.6 & &= 1.6 + 0.6 + 0.6 + 0.2 \\
 &= 3 & &= 3
 \end{aligned}$$

Therefore the algorithms correctly predicted the location as (3, 3).

## 4 Experimental analysis

### 4.1 Setup

For our tests, we utilised a corridor in Griffith University’s G09 building. Locations for Wi-Fi scans were chosen at three-metre intervals along the corridor. These are signified by the coloured dots in figure 2. During the tests, a total of

21 access points were discovered. We later filtered this list to access points with a specific network name (SSID), resulting in a collection of 4 access points. The locations of access points in the environment were not known.

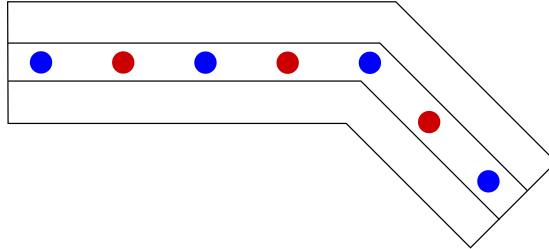


Figure 2: G09 building. Blue dots represent the training set, with blue and red dots being used for testing.

To prepare our dataset for processing we created a subset containing all of the data from every second location. This dataset, as represented by the blue dots in figure 2, contains four locations and was used as a training set. This dataset would be representative of one taking Wi-Fi readings along the corridor at intervals of 6 metres. All locations (i.e. both red and blue dots in figure 2) were used for testing the accuracy.

A test set containing unseen samples, such as that used in our experiments, is representative of a real-world situation in which a user may be standing in the middle of two training locations. Traditional approaches would pick one of the two locations, and could provide vastly different accuracy rates. The proposed MDWI approach allows for a location in the middle to be selected.

## 4.2 Algorithms

We ran our real world datasets through both a Bayesian Network (BN) and a K-Nearest Neighbour (KNN) algorithm. These two machine learning algorithms are commonly used in Wi-Fi localisation literature.[7] For both of the algorithms, we apply MDWI with  $n$  set to 5.

BNs already calculate probabilities as part of their classification process. We were therefore able to use these probabilities along with the coordinates of the locations in our weighted interpolation calculations. We used the University of Waikato’s Weka implementation of a Bayesian Network. The settings in version 3.6.9 of the software were left as defaults.

Unlike a BN, KNN algorithms do not calculate probabilities but rather output the sum of differences between two instances. As a smaller sum is a closer match, we took the inverse of the lowest-cost (or best-matching) for each location and calculated probabilities.

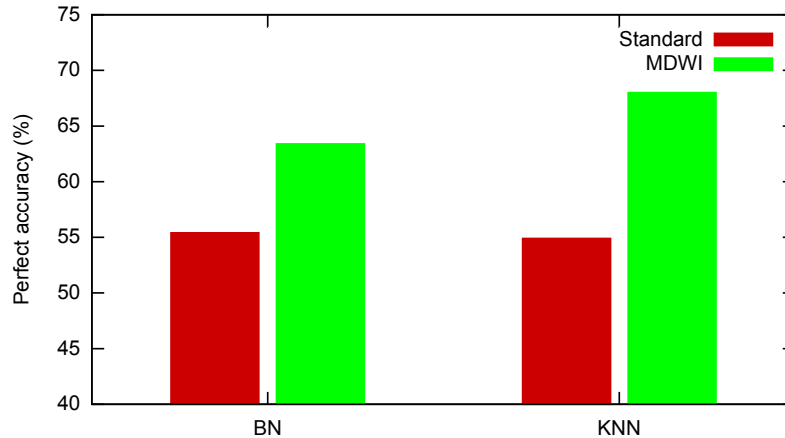


Figure 3: Perfect accuracy rate comparison

### 4.3 Results

The results from our tests are summarised in figure 3 and table 2. The figure shows the perfect accuracy rates of various algorithms, both before and after the application of MDWI. We define a perfectly accurate case as one which returns the same location as where the samples were originally taken. When MDWI is applied, we round the  $x$  and  $y$  coordinates to the nearest three metres, to match the granularity of the original readings. The figure shows perfect accuracy rates increase in all cases when MDWI is applied.

The percentages in table 2 are calculated by taking the actual location and comparing it to the algorithm’s prediction. Therefore a 0m error distance indicates the algorithm correctly predicted the location, 3m indicates an adjacent location, and so on. Again, the coordinates output from MDWI have been rounded to the nearest three metres.

The tests performed on the BN showed positive results. In these tests, MDWI shifted the majority of the error distribution towards smaller error distances, capping all errors at 6m. The BN with MDWI correctly predicted perfect matches 63.4% of the time, an increase of 8 percentage points. The number of 6m errors did slightly increase, however, we feel the 14% increase in perfect predictions offsets this.

The KNN tests also provided positive results. When KNN was used along with MDWI perfect accuracy was at 68%, an increase of 13.1 percentage points, providing the best result in all of our tests. However, MDWI’s application to KNN also had the unintended result of stretching the error distribution out to 15m. The main reason why this happened with KNN and not with the BN to the same extent was the distribution of the probabilities calculated for KNN were more evenly distributed than that of the BN.

The application of MDWI clearly provides a number of benefits. Accuracy

Table 2: Error distribution

Settings	Rounded Error Distance					
	0m	3m	6m	9m	12m	15m
BN	55.4%	42.9%	1.7%	–	–	–
BN+MDWI	63.4%	32.0%	4.6%	–	–	–
KNN	54.9%	42.3%	2.3%	0.6%	–	–
KNN+MDWI	68.0%	26.3%	3.4%	0.6%	1.1%	0.6%

rates can be increased when unseen intermediary locations are input as the algorithm will not immediately pick the closest-matching location. Additionally, if one is limited in the time spent and/or in the associated costs in training a system, MDWI can provide greater precision.

## 5 Conclusions and future work

In this paper we propose Multi-Directional Weighted Interpolation (MDWI) for any Wi-Fi localisation system, whether it be designed for human or robot navigation. MDWI utilises probabilistic-based weighting mechanism to output locations which have not been included in a training set. Our results show that, as with any localisation system, it is necessary to find a balance between the number of training locations and the desired level of final precision. The number of training locations are proportional to distance-based accuracy rates, however, more training locations also come at a higher cost. Our results show MDWI facilitates the ability for training costs to be halved without affecting accuracy rates to the same degree.

Future work will focus on utilising the theoretical concepts proposed in this paper to evaluate its performance in larger areas and areas which feature more dense training locations, as well as how signal strength variations between multiple devices are handled.

## References

- [1] P. Bahl and V.N. Padmanabhan. RADAR: an in-building RF-based user location and tracking system. In *Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 2, pages 775–784 vol.2, 2000.
- [2] D. C. Dimitrova, I. Alyafawi, and T. Braun. Experimental Comparison of Bluetooth and WiFi Signal Propagation for Indoor Localisation. In *Wired/Wireless Internet Communication*, volume 7277 of *Lecture Notes in Computer Science*, pages 126–137. Springer Berlin Heidelberg, 2012.
- [3] E. Elnahrawy, X. Li, and R. P. Martin. The limits of localization using signal strength: a comparative study. In *First Annual IEEE Communications*

*Society Conference on Sensor and Ad Hoc Communications and Networks*, pages 406–414, 2004.

- [4] Google Scholar. Bahl: RADAR: an in-building RF-based user location and tracking system. <http://scholar.google.com/scholar?cites=5301100132960192733>, August 2013. Accessed: August 13, 2013.
- [5] A. K. M. M. Hossain, H. N. Van, Y. Jin, and W-S. Soh. Indoor localization using multiple wireless technologies. In *IEEE International Conference on Mobile Adhoc and Sensor Systems*, pages 1–8, 2007.
- [6] K. Kaemarungsi, R. Ranron, and P. Pongsoon. Study of received signal strength indication in ZigBee location cluster for indoor localization. In *10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, pages 1–6, 2013.
- [7] M.B. Kjaergaard. A taxonomy for radio location fingerprinting. In *Proceedings of the 3rd International Conference on Location-and Context-Awareness, LoCA'07*, pages 139–156, Berlin, Heidelberg, 2007. Springer-Verlag.
- [8] B. Li, J. Salter, A. G. Dempster, and C. Rizos. Indoor positioning techniques based on wireless lan. In *First IEEE International Conference on Wireless Broadband and Ultra Wideband Communications*, 2006.
- [9] M. Ocana, L.M. Bergasa, M.A. Sotelo, J. Nuevo, and R. Flores. Indoor robot localization system using wifi signal measure and minimizing calibration effort. In *Proceedings of the IEEE International Symposium on Industrial Electronics*, volume 4, pages 1545–1550, 2005.
- [10] R. Want, A. Hopper, V. Falcão, and J. Gibbons. The Active Badge Location System. *ACM Transactions on Information Systems*, 10(1):91–102, January 1992.
- [11] M. A. Youssef, A. Agrawala, A. U. Shankar, and S. H. Noh. A probabilistic clustering-based indoor location determination system. Technical report, Department of Computer Science and UMIACS, University of Maryland, College Park, MD 20742, March 2002.
- [12] J. Zhou and J. Shi. RFID localization algorithms and applications—a review. *Journal of Intelligent Manufacturing*, 20(6):695–707, 2009.