Lessons Learnt From Designing Indoor Positioning System Using 868 MHz Radios and Neural Networks

Michał Meina
Faculty of Mathematics, Informatics and Mechanics
University of Warsaw
Warsaw, Poland
mail:mich@mimuw.edu.pl

Bartosz Celmer
Section of Computer Science
The Main School of Fire Service
Warsaw, Poland
mail:bart.celmer@gmail.com

Krzysztof Rykaczewski
Faculty of Mathematics, Informatics and Mechanics
University of Warsaw
Warsaw, Poland
mail:k.rykaczewski@mimuw.edu.pl

Abstract—This paper summarizes our approach and experimental evaluation of infrastructure-based Indoor Positioning System (IPS) designed to be used by First Responders. We are using 868 MHz single channel, power-efficient radio markers and RSSI (Receiver Signal Strength Indicator) fingerprinting. Artificial Neural Network translates vectors of RSSI constructed using mobile units into position. Special preprocessing needs to be applied to on-line signal to construct a vector for classification.

I. INTRODUCTION

Positioning system in GPS-denied environments (such as large buildings, tunnels etc.) can play crucial role in the Search and Rescue operations. Indoor Positioning Systems (IPSs) can improve not only safety of First Responders, e.g. Smoke Divers who have to search the building and safely withdraw having very limited oxygen supply, but also system like that could assist in decision making and risk management at rescue scene [1] by enhancing situation awareness of the Incident Commander. Yet, reliable information about deployment of resources (both humans and equipment) in the dynamical changing, decision-demanding environment is very challenging. This paper summarizes our approach for infrastructure-based indoor localization designed to be a part of risk management system for Incidents Commanders.

IPS can significantly optimize performance of fire brigade at incident scene in various aspects. Firstly, communication can be significantly enhanced as it was discussed in [4]—basing on interviews with experts and some on-field experiments the number of voice communication could be significantly reduced. Secondly, one of the most significant factor that is known to be source of accident (or near-miss incidents) at a fire scene is the lack of situational awareness1. Information about deployment of personnel is a key-factor at incident place. At the same time, it should be advised how to present the information. Amount of information generated at the scene can easily overwhelm [1] the Incident Commander.

Most reliable IPSs involve usage of infrastructure-based techniques, in which various transmitters (or beacons) are deployed in buildings beforehand. That enables us to position the receiver node carried by the subject. Different techniques can be used for such setup (see survey [2] for extensive summary). Following this survey, we have chosen to exploit radio Received Signal Strength Indicator (RSSI) fingerprinting, because of two main reasons: (1) relatively cheap and easy to deploy, and (2) accuracy can be easily tuned by adjusting number of anchor nodes.

The research was supported by the Polish National Centre for Research and Development (NCBiR)—Grant No. O ROB/0010/03/001 in the frame of Defence and Security Programmes and Projects: “Modern engineering tools for decision support for commanders of the State Fire Service of Poland during Fire&Rescue operations in the buildings”.

the algorithm needs to be fault-tolerant: damage of a single radio should not increase significantly the positioning error. Secondly, it is known that fire environment could alter signal propagation significantly [3].

The contributions of this paper are as follows:

- experimental evaluation of IPS based on RSSI fingerprinting and Artificial Neural Networks,
- discussion on hardware design of radio markers (anchor nodes).

II. RELATED WORK

Due to the growing demand for indoor positioning systems, wireless location is an important area of research in recent years. Many studies have been published concerning different types of localization techniques.

Both Personal Dead Reckoning (PDR) and Foot-mounted Pedestrian Navigation System use Inertial Measurement Units (IMUs) for path estimation. However, since the IMU position error accumulate during the procedure of walking, a lot of attention is paid to systems based on a pre-installed infrastructure. Nowadays, the most frequently used technology is based on radios, e.g.: pseudo-satellite transmitters, Radio Frequency Identification (RFID) markers and Ultra-Wideband (UWB) radars. This has many advantages, radio is not limited by the line-of-sight condition as radio signals can penetrate walls and diffract around objects [5].

Several methods have been proposed to estimate the location using sensor networks. Usually, the approach is based on reference nodes (beacons, anchor nodes), which positions are known. The position of the receiver is calculated from the information it receives from the beacons.

The position can be derived from distance estimates between the beacons and receiver node. Most radio receivers in a wireless system have the ability to measure the Radio Signal Strength Indicator (RSSI). This can be later translated to a distance by using a path loss model. Generally, the relation between RSSI and distance is determined by the following formula:

\[ \text{RSSI}(d) = P - R - 10\alpha \log_{10} d, \]  

where \( P \) is the transmitted power, \( \alpha \) is the path loss exponent which falls linearly and \( R \) is a constant that depends on the conditions of the environment, \( d \) is the distance from transmitting end [6].

Generally, three main methods are used for the problem of localization: trilateration, multilateration and fingerprinting. Trilateration and multilateration are based on the propagation model, conceptually simpler, but difficult to calculate in a complex environment—firstly distance needs to be estimated accurately, which involves usage of more expensive transceivers (e.g. UWB radio that uses ToF model2) and, secondly, environment (and its changes over time) modelling can be challenging. In contrast, fingerprinting is an empirical method in which signal attenuation in the building are measured and, therefore neither the signal propagation model is not used nor the building plan does not to be known.

Trilateration technique uses properties of triangles to determine the location, therefore it usually requires at least three access points on the surface. While using this technique precise distance needs to be measured (which usually is not achievable using RSSI). Precise distance is measured using different physical techniques: Angle of Arrival (AoA) is a method that locates the user by measuring the angle of incoming signal, Time of Arrival (ToA) is a technique based on the Time of Flight (ToF). While using this method the clocks of all physical units must be precisely synchronized and clock drift compensated. All this makes the final system more complex.

Multilateration is a navigation technique based on measurement of the difference between the distances to two or more stations located at known locations that transmit a signal in the indicated times [7]. It differs from trilateration in that it does not use absolute measurement of Time of Flight, but its differences (TDoA, Time Difference of Arrival). Position is then estimated by the intersection of hyperboloids which are places consisting of points having equal TDoAs. In this case, the problem can be represented as an optimization problem and solved using, for example, the method of the least squares or gradient descent method.

Due to the fact that certain signals can be disturbed by presence of obstacles, some extensions (like, e.g., multiwalls model) to above-mentioned methods were introduced.

Fingerprinting is another method frequently used in indoor positioning. In this technique radio signal strength is measured at different locations beforehand. During the first (training or off-line) stage signal strength data is collected is the physical location (usually up to 50 × 50 m²) to the training/labelling database, or to a non-linear mapping. In the second (on-line) stage of the mobile unit measures RSSI and compares its value to values held in the database. In result location with similar matching is returned. The location of the fingerprint technique requires an adequate number of reference devices and stable environment before calibration, because the result is sensitive to environmental changes, such as moving objects in a building that may have an impact on the properties of the signal. Fingerprinting can obtain good performance, since the noise arising from all obstacles is already included in the map. Therefore, we do not have to add to it any additional model.

The widely used basic matching algorithm used in fingerprinting is the k-Nearest Neighbour (k-NN) [8]. In the online positioning step the k-NN algorithm is searching for k neighbour closest (in the sense of the Euclidean distance) between classes of fingerprint database and the real-time RSSI values to determine the location. The Support Vector Regression (SVR) [9] as well as, Artificial Neural Network (ANN) are in widespread use as well [10], [11], [12], [13]. Comparison of different architectures of neural networks can be found in [12]. Moreover, there are attempts to combine such estimation with dead-reckoning navigation using foot-strapped inertial measurement units [14].
III. HARDWARE SET UP

We use transceivers based on RFM22 Hope Microelectronics co. silicon with 17 cm copper wire antenna (half-wave long). RFM22 have ability to work in very different modes (modulation, frequency, transmit power). Additionally, our setup involves computations of neural network on mobile device (Odroid-U3), sending it via ZigBee 867 Mbps (2.4 GHz) and displaying actual positioning using Recon Jet head-mounted display\(^3\).

The relatively low frequency (868 MHz) was chosen because of the high penetration ability of signal comparing to power consumption. Also, the noise from different devices operating in this particular frequency band is expected to be smaller (comparing to e.g. 2.4 GHz). It is worth to note that higher penetration gives us possibility to build sparse node networks, which drastically lowers the cost of overall system. Another issue is signal attenuation in smoke and fire environment which is known issue (see [3]) but there is not enough comparable results to choose the best operating band for this purpose.

Two type of radio nodes was constructed: anchor node (see Fig. 1a) and receiver node (Fig. 1b). Anchors send periodically small portions of data including identifier and message number. Receiver Node is gathering those packages while establishing RSSI and reports it to processing unit.

Configuration of RFM22 was as follows: 868 MHz frequency transmission band, FSK modulation without Manchester encoding (error detection technique—disabled for shorter time of transmission) and +17 dBm mode (transmit power). During initial test we confirmed sufficient wall penetration and expected RSS loss.

Every anchor node operates on exactly the same frequency and, because of that, two radios which transmit their signal can drown each other. Two or more radios that are in mutual coverage area cannot transmit their data in the same time. Therefore, transmission synchronization needs to be performed to overcome problem of mutual jamming. Nevertheless, direct clock synchronization is very complicated in Wireless Sensor Networks (see [15] for overview of the problem), especially in indoor environment (where GPS-based synchronization is unavailable).

Straightforward, node-independent mutual jamming prevention technique based on randomized transmission was implemented. Node number \(i\) transmits its mark which usually lasts for about 15 ms, and then radio goes into sleep mode for \(T_i\) ms. Idle time is picked randomly after each transmission from the interval

\[
t_{\text{w min}} < T_i < t_{\text{w max}}.
\]

This way, idle time is long enough to allow other radios to transmit their data and short enough to retransmit packet, if it was dropped while mobile is not moved far. Due to the high noise in RSSI estimation it is important to get as many readings as possible. The data is later preprocessed using moving window technique (see Section V-A).

Figure 2 shows simulation result for selected values of \(t_{\text{w min}}\) and \(t_{\text{w max}}\) with regard of the number of anchor nodes. Increased number of anchor nodes obviously lead to increased sampling rate at Receiver Node. This, however, can be done only to some extent, after which sampling rate is degrading (because of the collisions in transmission). Peaks at Fig. 2 indicates more or less “optimal setup”. Having estimated number of nodes that can jam each other on deployed building interval of sleeping time should be adjusted using this simulation.

Sampling rate estimated from empirical data (which was 16.3 ± 2.5 rps) in our experiment is higher than expected (depicted by the box plot on the right-hand side of Fig. 2). In our experimental deployment (see Fig. 6) distant nodes are on the edge communication reach and, therefore, they have limited possibility to drown each other.

IV. SYSTEM DESCRIPTION

Overall system processing schema is illustrated in Fig. 3. Operation of IPS that is based on fingerprinting is divided into two phases: off-line and on-line. In the first fingerprints are collected and learning procedure is performed. In on-line phase, on the other hand, incoming signal is processed “on-the-fly” and algorithm outputs position.

At step 1 fingerprints are collected (numbers in circles refers to Figure 3). Radio signal strength is recorded using

\(^3\)http://www.reconinstruments.com/products/jet/
mobile device at predefined places (with known GPS position) and used to construct a feature vector (fingerprint) for particular point.

In the on-line situation at the beginning (3) RSSI signal needs to be preprocessed to create a test vector for ANN. Due to the fact that RSSI values from different radios did not arrive at the same time, sliding window technique is exploited—due to the small sampling rate fast movements of mobile unit can degrade precision of test vector formulation (windowing technique introduces lags). Later we remap entries of this vector to the interval \([0,1]\) in order to use the neural network at step 2.

Node 4 is used to resolve the method that will be used for positioning. It turns out that the simple thresholding of signals is inadequate, because some radios can be transmit with lower RSSI value outside than inside. However, it is true that the signal strength with decrease with the distance from the building. Indoor/outdoor detection is beyond the scope of this paper.

Classification (mapping RSSI to position) is performed at step 3. Just after it simple mapping to GPS coordinates is performed.

At the end 4 path needs to be post-processed (expected high RSSI noise introduces a lot of distortions on path). Kalman Filter is a good candidate for solving this problem, because of the motion model that can be expressed by it.

V. NEURAL NETWORK-BASED FINGERPRINTING

This section describes our framework for IPS based on Artificial Neural Network.

Most commonly used type of Artificial Neural Network (ANN) consists of several layers: the input layer is connected to layers of hidden units, which provide information to the output layer. For learning ANN the most commonly used method is the backpropagation algorithm. It tries to adjust weights of each neuron in order to reduce the error between the desired and calculated output. In this way, the neural network learns how to map input to output. The aim of the network is not only to restore the training input data but also to generalize the data to new situations (by interpolating capabilities). The number of input nodes and hidden layers depends on the design issue and depends on the number of base stations deployed in the environment.

A. Input data preparation and training set construction

The data collection process involves marking of the reference points on the floor and making measurements for a 30–60 s. All points laid on the same plane. In this way at every point we received a number of recordings consisting of RSSI which comes from different radios. Therefore, we obtained a RSSI fingerprint log consisted of triplets

\[
i, \rho_i^{jk}, P_j,
\]

where \(i = 1, \ldots, N,\) denotes radio number, \(P_j = (\text{latitude, longitude, elevation})\) is a position of \(j\)th fingerprint and \(\rho_i^{jk}\) is a \(k\)th RSSI value recorded at point \(P_j,\)

Since all radios may not be visible at once (due to interference and momentary jamming of radios), in order to obtain a vector of RSSI signals from all radios (which will be used as an input to network) we had to aggregate recordings at a given point. Therefore, as input corresponding to point \(P_j\) and reading \(k\) we took set

\[
\{ \text{avg}(\{\rho_i^{jk}\}_{i=1}^{K_j/2}) \mid M \text{ rand } \{k_i^{jk}\}_{i=1}^{K_j/2} \subset \{1, \ldots, K_j\} \}
\]

(4)
i.e. we averaged random subsamples of recordings (by taking half of the recordings) for point \(P_j.\) In order to get rid of the noise we do it \(M\) times. Therefore, \(M\) can be interpreted as an aggregation (folding) parameter.

Let us define point signals in time \(t\) from \(N\) radios in \(j\)th point by the following

\[
\rho^j(t) = [\rho_1^j(t), \ldots, \rho_N^j(t)].
\]

(5)
Such signals consist of RSSI recordings aggregated and averaged as above.

Due to issues described below, which are related to neural networks, we need to map GPS coordinates using affine scaling into interval \([0,1]\). Similarly, RSSI values from radio are converted into \([0,1]\).

B. Network architecture

Neural Network that we use for IPS is depicted in Fig. 4. It consists of input layer constructed from RSSI vector for actual reading, output layer denotes position in three dimensional space (latitude, longitude and elevation) and the \(L\) number of hidden layers. We have used sigmoid activation function, therefore units in input and output vectors need to be mapped by scaling into interval \([0,1]\).

The number of hidden layers and number of neurons at each layer should be tuned accordingly to specific task. Nevertheless, it is important to say that this particular architecture depends significantly on number of anchor nodes used for IPS system. It is hard to estimate time complexity of the learning algorithm. Estimation for the worst case scenario is

\[
M N L \prod_{i=0}^{L} \# W_i.
\]

(6)

Fig. 4: Neural network setup.
where \( \#W^i \) is the number of weights in layer \( i \) and \( L \) is the number of layers in the network. We see that it depends strictly on the size of the building, since additional reference nodes need to be added.

Learning takes quite long, depending on the permissible error. For example, for our problem with permissible error equal to 0.01 and folding parameter equal to 20 (which means that the number of samples grows by the factor of 20) it took between 4 and 10 h on a single core. The number of iterations was set to 10000. It is important to notice that quality of network classification depends on the number of iterations and goal set.

C. On-line signal classification

Due to the possible interference and temporary signal deficiencies input data in on-line classification need to be preprocessed first. During the movement the mobile unit can receive readings from different anchors in different times (radio reading are sparse and unevenly sampled), therefore we used sliding window technique to receive feature vector at given time.

In order to impute missing values of signal strength we can use local linear approximation as follows

\[
\rho(t + \Delta t) = \rho(t) + \rho'(t)\Delta t + \mu
\]

\[
= \rho(t) + \beta d'(t) \log_{10} d(t) + \mu,
\]

where \( \beta \) is some constant and \( \mu \) denotes higher order terms with small magnitude. Assuming that locally velocity of the subject is constant (\( d'(t) = \tau_{loc} = \text{const} \)) we use it to smooth the signal. The result is shown in Fig. 10 and was not so impressive as supposed. Therefore, we used linear regression.

Given fingerprints as in (3) we want to obtain input signal for the network as in (5).

Since at given time not all radios may be visible we define moving window \( W \) of length \( \tau \) for time series \( s \) (till time \( t \)) as

\[
W_{\tau}(s)(t) := [s(t - \tau), \ldots, s(t)]
\]

which will collect the signal strength in a short period.

To this end, we performed moving average on RSSI from the last two seconds. If there was no radio signal from a given radio in the given window, then we put 0. In short,

\[
\hat{\rho}(t) := \left[ \text{MA}_k(W_{\tau}(\rho_1)(t)), \ldots, \text{MA}_k(W_{\tau}(\rho_N)(t)) \right],
\]

where \( \text{MA}_k \) stands for moving average operator of length \( k \). Signal prepared in this way can be used as an input to the neural network. For post-processing we used the Kalman filter. Therefore, path are smoother and better corresponds to reality.

D. Network architecture tuning

We found out that standard heuristics, like taking two hidden layers with number of neurons in the second layer being half of that in the first layer (see [16] for more information) works quite fine. We used very large first layer since we need embed data in high dimension and obtain overfitting in order to discriminate it. Next layer is smaller but we get better generalization properties and avoid overfitting. Such architecture can be obtained when analysing in detail Fig. 7.

VI. EXPERIMENTAL SECTION

We conducted experiment using 18 anchor nodes deploying them on the one floor of approximately 30 \( \times \) 30 m building. Fig. 6 shows displacement of radios and fingerprints. Note that the data was collected only in selected rooms and passages. Additionally, the figure shows bicubic interpolation of RSSI signal strength for 5th radio basing on fingerprint features.

We performed collection at 119 points gathering 5991 records (samples) overall. Each position was recorder for about 18.1 \( \pm \) 4.2 seconds which gives us 297.3 \( \pm \) 90.2 records on average. Points were not distributed uniformly, but they were chosen in such a way that there were no large area without samples. Moreover, significant error was introduced while we manually collected GPS positions of reference points.

Two experiments were conducted: stationary position estimation and movement/path reconstruction. For training the neural networks we have used folding parameter \( M = 20 \). It may seem rather small, but that enables us to perform parameter sweeping through different network architectures in reasonable time (see Fig. 7).
A. Position in stationary points

In the first experiment we tested position estimated at stationary points (without movement). Recorded fingerprints were divided into two separate sets in ratio 70/30. Points was chosen manually and are depicted in Fig. 6. Test sample was prepared according to Equation (4).

As it can be seen, our heuristics works quite good. For example, when there is about 10–14 neurons in the first layer the average error falls below 0.5 m. Larger networks do not guarantee an improvement of the results (even though some of them were significantly better).

B. Paths

Second experiment was focused on path reconstruction. We asked a subject to perform a walk-through the building with receiver node. Results are depicted in Fig. 8–10. True path was marked on figures using video camera recordings.

As it was already written, most of the errors in the motion is introduced by the low frequency of refreshing rate and, thus, there is the necessity to use of sliding windows.

In smaller networks we can see that path is being pulled off (Fig. 9).

Another issue is the necessity of usage of sliding window technique—it introduces a certain lag into the feature vector construction. Simply speaking, different RSSIs are recorded at different places therefore we do not know signal strength precisely in time. The problem is illustrated on both previous mentioned figures at the path in lower-right room.

Fig. 7: Position estimation error using ANN with \( \#W^1 \) and \( \#W^2 \) neurons on the first and the second layer respectively (lower figure). Upper figure shows column-wise summed up results for \( H^1 \) layer.

As it can be seen, our heuristics works quite good. For example, when there is about 10–14 neurons in the first layer the average error falls below 0.5 m. Larger networks do not guarantee an improvement of the results (even though some of them were significantly better).

B. Paths

Second experiment was focused on path reconstruction. We asked a subject to perform a walk-through the building with receiver node. Results are depicted in Fig. 8–10. True path was marked on figures using video camera recordings.

As it was already written, most of the errors in the motion is introduced by the low frequency of refreshing rate and, thus, there is the necessity to use of sliding windows.

In smaller networks we can see that path is being pulled off (Fig. 9).

Another issue is the necessity of usage of sliding window technique—it introduces a certain lag into the feature vector construction. Simply speaking, different RSSIs are recorded at different places therefore we do not know signal strength precisely in time. The problem is illustrated on both previous mentioned figures at the path in lower-right room.

Our initial idea was to use forecasting on RSSI signal that can minimize lagging issues with standard rolling window techniques. Ability of forecast the signal strength lies on the assumption that RSSI is dependent on motion (acceleration and decelerations are not random), see Equation (7). Therefore, we tested the idea using local linear regression instead of rolling mean. Results are depicted in Fig. 10 and show clearly that this simple forecasting is not working due to the large impact of noise to path estimation. Nevertheless, the idea of compensating the signal sampling with motion prediction is worth pursuing in the future works.

C. Discussion

Fingerprint coverage does not have to be dense (neural networks have very good interpolation capabilities), but we noticed a problem with the estimation of places “on edges,”
where the paths are pulled towards more places with more dense fingerprint coverage. Some special case should be applied in order to overcome that.

It is worth noting that more anchor nodes not necessarily means better performance—sparse node deployment allow higher signal sampling rate and, therefore, allow better estimation of high-velocity motions. Observe, however, that we cannot assume that radios have different sampling, because we do not know whether the situation when radio sees only few neighbours will change in the future. During fire some wall may broke down and opens new way for the signal.

We also noticed a negative correlation between the velocity of movement of the subject and location accuracy. This is obviously due to the fact that we get only a few samples of RSSIs for a given position. In result the path can oscillate around the correct location.

Moreover, it can be observed that iron oxygen cylinder carried by the firefighters can distort the signal very badly. For example, there were situations where orientation (rotation along the axis) lead to tremendous improvement of the position.

REFERENCES


