Low Complexity Wireless Sensor System for Partial Discharge Localisation

Ephraim T. Iorkyase[1], Christos Tachtatzis[1], Pavlos Lazaridis[2], Ian. A. Glover[2], Robert. C. Atkinson[1]

[1] Department of Electronic and Electrical Engineering, University of Strathclyde, Royal College Building, 204 George Street, Glasgow, G1 1XW, UK
[2] Department of Engineering and Technology, University of Huddersfield, HD1 3DH, Huddersfield, UK
*ephraim.iorkyase@strath.ac.uk

Abstract: This paper describes a key element of any modern wireless sensor system: data processing. We describe a system consisting of a wireless sensor network and algorithmic software for condition-based monitoring of electrical plant in a live substation. Specifically, the aim is to monitor for the presence of partial discharge using a matrix of inexpensive radio sensors with limited processing capability. A low-complexity fingerprinting technique is proposed, given that the sensor nodes to be deployed will be highly constrained in terms of processing power, memory and battery life. Two variants of artificial neural network (ANN) learning models (multilayer perceptron and generalised regression neural network) that use regression as a form of function approximation are developed and their performance compared to k-nearest neighbour and weighted k-nearest neighbour models. The results indicate that the ANN models yield superior performance in terms of robustness against noise and may be particularly suited for PD localisation.

1. Introduction

Wireless Sensor Networks (WSN) and their associated systems are now mature technologies. They provide the capability for both remote and distributed sensing operations in applications as diverse as human health monitoring [1] [2], animal health monitoring [3], turbine monitoring [4], and condition-based monitoring [5].

In this paper we describe the use of a low cost WSN for condition-based monitoring of electrical substation equipment. As such this represents a highly industrially focused application. The overall wireless sensor system (WSS) developed, consists of a WSN (including customised hardware) and the necessary software and data processing components.

Specifically, the WSS being developed is designed to monitor for the presence of partial discharge (PD) which is a well-known forerunner to asset failure in substations. It is caused by localized insulation defects such as the existence of voids, and other impurities in an electrical insulation system be it solid, liquid or gas [6] [7]. Early detection of PD permits preventative maintenance to be employed and/or fault diagnose and repair avoiding catastrophic failure which may result if the asset is not serviced. Timely intervention reduces the costs attributable to unplanned outages and permits conditioned-based maintenance.

In order to determine which asset within a substation is exhibiting PD, its location must be estimated. Given that the discharges themselves give rise to electromagnetic pulses, they can be detected using appropriate radios. Thus in this application, we use the WSN radios as sensors. In fact, radiolocation of an electromagnetic source can be accomplished using a variety of techniques that often include time difference of arrival (TDoA), angle of arrival (AoA), or received signal strength (RSS) [8] [9] [10] [11] [12] [13]. It has been established that these parameters (TDoA, AoA and RSS) are location dependent and can be used to locate PD sources.

TDoA and AoA methods are computationally expensive and have energy-hungry signal processing costs. RSS-based approaches are more appropriate to low-complexity implementations but at the cost of reduced localization accuracy [14] [15]. The reduced accuracy of purely RSS-based methods is attributable to their susceptibility to interference due to multipath propagation, path loss attenuation and/or signal shadowing. Approximate RF propagation models from literature are not suitable for PD source location due to the complexity of real-life radio environments in which PD occurs. This motivates an investigation into the feasibility of using machine learning based fingerprinting technique for PD localization. This method not only obviates the need for estimating radio channel parameters but also uses multipath to its advantage to provide robust PD localization. It learns the unique spatial signatures created by PD signal strength to estimate PD locations.

Our low cost/complexity solution is to deploy a WSN based on off-the-shelf commercial radio nodes to monitor PD radio emissions with some bespoke hardware [16] [17]. The wireless nodes will be deployed in an approximately regular grid topography across the substation. The inexpensive solution proposed permits the deployment of a permanent substation-wide continuous monitoring system in real-time. With this solution, the network built autonomously a fingerprint map with the following process: each sensor node emits an emulated PD pulse of known transmission power in a round-robin fashion. Each RF emission is monitored by the other wireless nodes. This permits each node to build a path loss map of the substation from its unique perspective. Real PD emissions will have an unknown transmission power, and

this motivates the use of relative (ratios) RSS between nodes to define a fingerprint. When true PD occurs, the ratio of RSS at the various nodes can be used to infer location via machine learning based interpolation. Thus, our WSN and associated system components perform distributed sensing. High location accuracy is required if the system is to be effective. Accordingly applying a simple path loss model to infer location is not appropriate, rather sophisticated machine learning algorithms are required. It is this latter aspect of a WSS that is the prime focus of this paper.

An advantage of a software-based machine learning approach is that the model can be retrained to accommodate changes in the substation topology by rebuilding the fingerprint map as described above. This retraining could be initiated manually or periodically. Given that PD pulses are impulsive and have a duration in the order of picoseconds, the network could be retrained within seconds of a change in topography.

The rest of the paper is organized as follows. Section 2 presents the description of the PD localization problem. Section 3 describes the experimental procedure. Section 4 describes the location fingerprinting technique. Section 5 provides the formulation of the localization algorithms. In section 6, experimental results are presented and discussed with conclusions drawn in section 7.

2. Problem Description

The key problem of PD localization is how to effectively estimate of the function:

$$f: R^M \rightarrow R^n \qquad (1)$$

which relates the mapping of PD fingerprints from *M* distinct sources onto their locations in n-dimensional coordinate space.

In this work, the PD localization problem is considered in a 2-dimensional space. The locations and characteristics of the PD pulses are not known a priori. The electromagnetic signals generated by the sensor nodes can be captured and recorded by appropriate wireless sensors positioned in the vicinity of the discharge source. The radio environment is modelled as a finite location space $L = \{l_1, \ldots, l_m\}$ of m discrete locations, where $l_i = (x_i, y_i)$ is the coordinate of a PD source.

Suppose r denotes an observed RSS ratios recorded by a sensor nodes placed in the area of interest to capture the RF signals. If there are P sensor nodes and M distinct PD sources, then the location fingerprint f is given as an $M \times P$ matrix:

$$f(l_i) = \begin{bmatrix} r_i^1(1) & \cdots & r_i^1(p) \\ \vdots & \cdots & \vdots \\ \vdots & \cdots & \vdots \\ \vdots & \cdots & \vdots \\ r_i^M(1) & \cdots & r_i^M(p) \end{bmatrix} \qquad (2)$$

The columns of the matrix in (2) represent RSS ratio vectors recorded by a sensor node from all the *M*, PD sources. The major challenge is to estimate as accurately as possible the PD location from location fingerprint (RSS ratios). The PD sources and receiving nodes are assumed to be static during measurement.

3. Experimental Procedure

3.1. Measurement Setup

In order to evaluate the feasibility of deploying a WSN sensor array for PD monitoring and localisation, PD data was collected in a 19.20 m × 8.40 m laboratory at the University of Strathclyde. The measurement campaign took place over a two-week period in the evenings so as to minimize the effects of people interruption. The environment is characterized by multipath propagation which is a result of cluttered objects including metallic ones. Fig. 1 shows the floor map of the laboratory with all the RF signal sampling points indicated by the red dots. In all there are 144 sampling points located on a uniform grid with spacing of 1m. A picosecond pulse generator was used to generate emulated PD traces at each of the sampling points, 20 pulses were generated at each location providing an overall sample set of 2880 pulses. As a proof of concept, three RF sensors positioned at predefined locations in the laboratory were connected to a multichannel digital oscilloscope to capture and record PD measurements. The experiments utilized commercial-off-the-shelf antennas: ¼ wavelength antennas operating at 173MHz. The true frequency response of the antenna is depicted in Fig. 2. The maximum gain of the antennas is in fact 200.7 MHz.

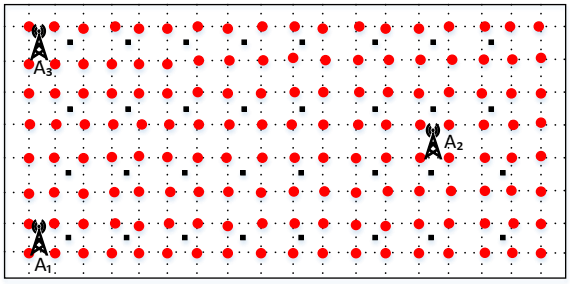


Fig. 1. *Measurement grid for measurement campaign*

Fig. 2. *Antennas frequency response*

Where deployed as part of an industrial solution, the sensors would be deployed as part of an Internet of Things platform consisting of onboard processing and a WirelessHART radio. Embryonic trials have already taken place to determine the applicability of this technology [5]. Fig. 3 shows samples of the captured PD traces. The PD data acquired from

measurement were sampled at 2 GS/s. This sampling rate allows the signals to be captured with high resolution.

3.2. Data Collection

For the purposes of modeling, a PD dataset was gathered from the measurement setup described above via the three sensors. At each location, 20 consecutive RF measurements were made, giving a total, 2880 RF samples. A second, independent, test set was gathered in the same way but at 32 distinct locations (the inter-grid black squares in Fig. 1). The first dataset was further divided into training and validation subsets, which were used to train and optimize the models as described in section 5.1. The independent test set was subsequently used to benchmark the performance of the models. The averages of the signal strength ratios of the RF samples collected were computed and used as fingerprint input vectors to the localization algorithms. Fig. 4 shows PD signal strength pattern in the radio environment for each of the three sensors. The figures reveal the unique spatial signatures created by PD signal strength at different locations which facilitate the application of learning models.

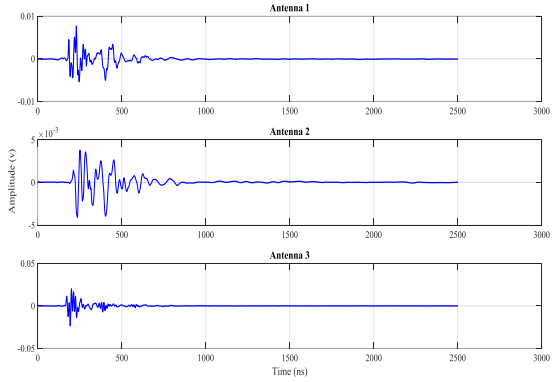


Fig. 3. *Recorded PD signals*

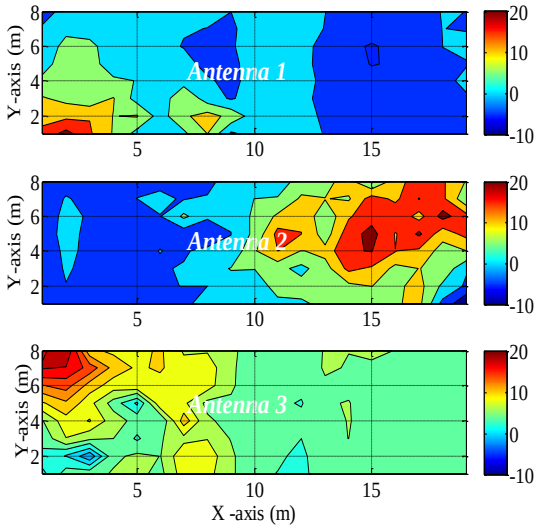


Fig. 4. *Spatial Pattern for antennas 1, 2, 3*

4. Location Fingerprinting

Location fingerprinting (LF) takes advantage of multipath propagation [18] since this contributes significantly in creating the needed spatial patterns for localization. The LF approach maps RF patterns (in this case RSS) to known spatial locations, then uses this mapping to infer previously unknown locations from RF patterns. The underlying assumption is that at every point in the propagation space PD exhibits distinct signal features.

The procedure is divided into two phases: a database construction phase and location estimation phase [19] [20] [21]. The database construction phase involves building a table of fingerprints (features) associated with a set of known locations using emulated PD pulses with known transmission power. Raw RSS measurements are frequently used as a fingerprint for radio beacons that are characterized by constant transmission powers [22]. However, given that the energy emitted by each PD event may be different due to progressive nature of PD severity as deterioration continues and the fact that different types of PD occur in nature, absolute RSS is not well suited to this problem. A more robust fingerprint is the ratio of RSS components received by the multiple receivers. In this paper, the signal strength ratios (SSR) between pairs of wireless sensor nodes are used as the location fingerprint.

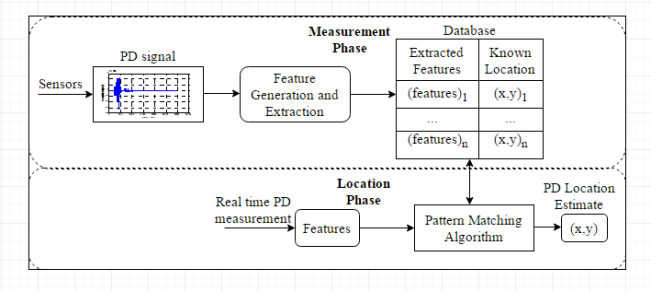


Fig. 5. *Location Fingerprinting System*

During the localization phase, the location of a PD source is estimated by comparing the fingerprints in the database with the real-time RSS. Fig. 5 shows the framework of the location fingerprinting system used for this application.

Assuming $P = \{p_1, \ldots, p_n\}$ is a set of nodes deployed in the substation and $L = \{l_1, \ldots, l_m\}$ represents the finite location space. Each location feature space, l_i can then be represented by a pair of nodes $p \in P$ and a measured signal strength value $r \in R$ where $R = \{r_1, \ldots, r_p\}$. The signal strength ratio (SSR) is defined at each location for a unique node pair $p_i \times p_j \in P \times P$ with the constraint $i < j$ for uniqueness. The SSR for receiver p_i and p_j can be computed for an observation measured at location $l = [(p_i, r_i), (p_j, r_j)]$ as follows;

$$SSR(p_i, p_j) = \frac{r_i}{r_j} \tag{3}$$

5. PD Localization Models

5.1. Model validation procedure

The overall goal of modelling is to make accurate predictions. In this work, the cross validation method is used for estimating the accuracy of PD localisation model's predictions on unseen cases. The optimised model is the one that makes the most accurate predictions. The general idea of cross validation is to divide the data sample into a number randomly drawn, disjointed sub-samples/segments otherwise known as v-folds. And for a fix value of any parameter, the model under consideration is applied to make predictions on the vth segment (i.e, use the $v-1$ segment as examples) and evaluate the error (mean squared error). This procedure is then applied successively to all possible choices of the parameter under consideration. At the end of the folds, the computed errors are averaged to yield a measure of stability of the model. These steps are then repeated for various values of the parameter and the value with the lowest error is then selected as the optimal value. This procedure is used to optimise the PD localisation models developed in this paper.

5.2. K Nearest Neighbour

K-nearest neighbour (KNN) [19] [23] is one of the best-known machine learning algorithms used for function approximation which smoothly interpolates between known samples. It is an intuitive fit with the problem of inferring the propagation environment between arrays of sensor nodes. In the context of PD localization, KNN regression consists of mapping RSS inputs onto dual output corresponding to PD location in 2-dimensional space. The underlying assumption is that all samples in a 'local' region within the location space have similar location fingerprints. In its simplest form, KNN computes the location of a real-time PD as the arithmetic average of the coordinates of *K* nearest neighbours. The nearest neighbours of any real-time PD RSS are determined by means of distance similarity metric in feature space. The most commonly used distance metric is the Euclidean distance [19]. Suppose there are N location fingerprints (SSR) in the database expressed as $R = \{r_1, \ldots, r_m\}$ and a real-time SSR vector, $S = (s_1, \ldots, s_p)^T$ is collected from p receiving nodes. The Euclidean distance between real-time sample s_i and fingerprint r_i in feature space is given as follows;

$$d_{fin}(S,R) = \sqrt{\sum_{i=1}^{p}(s_i - r_i)^2} \qquad (4)$$

Another distance metric that can be applied to the KNN regression algorithm is the Mahalanobis distance [24]. Given the location fingerprint (SSR) vector R in the database, a real-time SSR vector S and a covariance matrix Σ, the Mahalanobis distance is:

$$d^{m}_{fin}(S,R) = \sqrt{(S-R)^T \Sigma^{-1}(S-R)} \qquad (5)$$

The Mahalanobis distance is often preferred because it accounts for the variance of each feature and the covariance between features. KNN localisation of PD sources incudes database construction and the location estimation phases. Database construction comprises the collection and storage of known examples. Each example consists of a data point having PD fingerprint (i.e SSR from a source) labelled with its physical location coordinate. In the location estimation phase, KNN finds K examples in the database whose SSR values are closest (most similar) to the new PD observation in feature space (of fingerprint). In the nearest-neighbour calculations each fingerprint is represented as a vector of average SSR with entries for each receiver sensor. KNN prediction of PD location is based on averaging the physical location coordinates of the selected closest examples otherwise known as K-nearest neighbours. The estimated PD location is given by:

$$(\hat{x},\hat{y}) = \frac{1}{k}\sum_{i=1}^{k}(x_i, y_i) \qquad (6)$$

where $\hat{x}$ and $\hat{y}$ are the estimated coordinates. x_i, y_i are the coordinates of the kth nearest neighbour.

The value of the parameter *K* has an impact on the performance of the KNN localization model. The optimal value of *K* is obtained through Cross validation described in section 5.1.

The basic KNN regression described above allows all the *K* nearest neighbours in the feature space to contribute uniformly to PD location estimation regardless of their distances from the real-time sample in the coordinate space. The algorithm can be improved upon by using a scheme which weights each of the *K* neighbours such that nearer neighbours contribute more to the final location estimate than neighbours faraway. The weight of each neighbour is determined by taking the inverse square of the neighbour's similarity distance from the new observation. One advantage of the improved KNN algorithm is its ability to smooth out the impact of noisy training data by taking the weighted average of the nearest neighbours. This enhances localization accuracy. The improved algorithm, called the weighted *K*-nearest neighbour (WKNN), replaces (6) with:

$$(\hat{x},\hat{y}) = \sum_{i=1}^{K} w_i(x_i y_i) / \sum_{i=1}^{K} w_i \qquad (7)$$

where

$$w_i = \frac{1}{d^2(r,r_i)} \qquad (8)$$

Equation (8) is the weight of each neighbour and d is the fingerprint similarity distance.

Despite the advantages of the KNN algorithm such as simplicity of implementation and the facility to trade-off accuracy and computational complexity, issues have been revealed that can adversely affect its performance in PD localization. KNN is based on a smoothing approximation. In reality the locations in physical space do not perfectly correspond to locations in the fingerprint, or signal, space. Since the estimates are a function of the grid point's position localization accuracy is dependent on the resolution of the

grid points. Unfortunately, the PD signal measured by the sensor nodes represents a complex, nonlinear surface which cannot be captured by KNN algorithm. This has motivates the use of an artificial neural network (ANN) due to its ability to accommodate highly nonlinear relationships.

5.3. Artificial Neural Network

Artificial neural networks (ANNs) [25] are the most commonly used computational learning techniques for solving function approximation problems. They essentially imitate the learning process of biological neural networks through the use of interconnected nodes, often called neurons. In its basic form, a neuron computes its output using a weighted sum of its inputs and its activation function. The activation function introduces non-linearity and robustness. Each connecting link between nodes is associated with a weight that can be tuned based on experience, making the neural network adaptive and capable of learning. Given sufficient neurons and a set of input-output data, the ANN can be trained to approximate any continuous function arbitrarily well. The inherent noise immunity of ANNs and their ability to learn, and subsequently recognize complex and non-linear relationships between input and output vectors without any prior knowledge of the relationship make them attractive for these applications. Two variants of the ANN; the multilayer perceptron [26] [27] and generalized regression neural network [28] have been developed to estimate PD source locations given the received signal strength of the PD traces.

5.2.1 Multilayer Perceptron: The multilayer perceptron (MLP) is the most popular, and widely used, feedforward neural network [25] model for any kind of input-output mapping problem. In the context of PD localization, it consists of a nonlinear mapping of the PD RSS input onto the dual output variables representing the source location coordinates. The MLP network consists of an input layer, a hidden layer and an output layer as shown in Fig. 6.

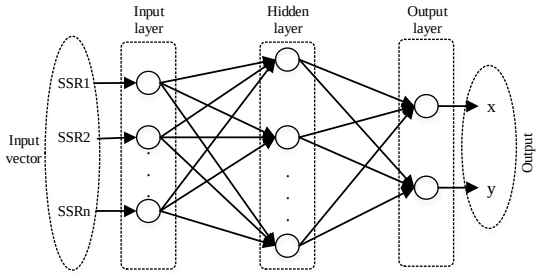


Fig. 6. *MLP model for PD localization*

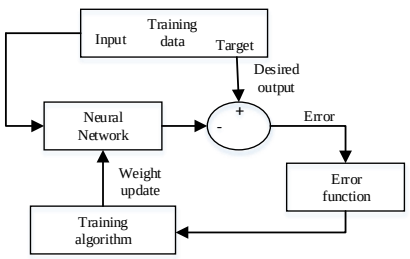


Fig. 7. *MLP training procedure*

A linear activation function in the input and output layers and a sigmoidal activation function is used in the hidden layer to provide robustness against extreme values. MLP is based on the back propagation training algorithm [26]. During the training phase, the network is trained to form a set of SSR values as a function of PD locations. The inputs to the network are the SSR values and network outputs are the corresponding location coordinates. Each sample is presented to the input and the error between the network outputs and the desired outputs is obtained. The neuron weights are then adjusted to minimize the error. The MLP training scheme is as shown in Fig. 7. This is an iterative non-linear optimization technique with an initialization stage. v-fold cross-validation is used to determine the optimal configuration of the network. The original training data is randomly divided into v equal-size subsets (the folds). In each case, one of the v subsets is used as validation data and the remaining are used for training. The cross-validation process is repeated v times. The average of the v results from the folds represents the test accuracy of that particular network. In this work a 10-fold cross-validation is used. From all the networks tested by cross-validation the feedforward 3-4-2 structure of the neural network with four neurons in the hidden layer has the best accuracy. The Bayesian Regularization algorithm is used to perform weight selection and optimization. The trained network is employed to simulate the test (unseen) data. The network uses the knowledge acquired during training to provide interpolated values for the location coordinates of the test data.

5.2.2 Generalized Regression Neural Network: Given that only a limited number of sensors will be deployed in the substation, then only a small training set will be available. This motivates the use of a probabilistic neural network, consequently a generalized regression neural network (GRNN) is proposed. This network uses a one-shot learning algorithm [28] [29] that, unlike MLP, does not require an iterative procedure for training. It is based on kernel regression and can approximate an arbitrary function. In this study, GRNN carries out a nonlinear mapping between the PD RSS and PD source location, drawing the function directly from the training data. It is consistent and its estimate always converges to a global minimum. The GRNN is related to the radial basis network and has a fixed structure with an input layer, a hidden (pattern) layer, a summation layer and an output layer that are fully connected as shown in Fig. 8.

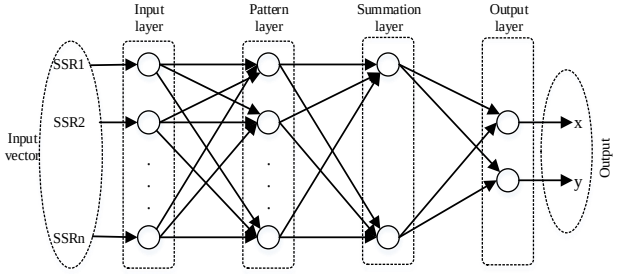


Fig. 8. *GRNN model for PD localization*

During training, the relationship between the fingerprints and the PD location coordinates is memorized and stored. In contrast to the MLP model, weight initialization is not

required. Once the inputs are presented to the GRNN the weight parameters are determine instantly. The weights between the input and each pattern neuron are set to unity. In each pattern unit, a radial (Gaussian) basis function is used as an activation function to calculate the output value such that the output from the pattern layer is given by

$$\theta = \exp[-\frac{(f-f_i)^T(f-f_i)}{2\sigma^2}] \quad (9)$$

where f and f_i are the PD fingerprints and σ is the spread or the smoothing parameter.

The summation layer is subdivided into S-summation neuron (summation units) and D-summation neuron (a single division unit). The S-summation neuron determines the sum of the weighted outputs of the pattern layer. The D-summation neuron determines the unweighted outputs from the pattern layer such that the estimated output to an unknown input vector is:

$$\hat{l} = \frac{\sum_{i=1}^{n} l_i \theta}{\sum_{i=1}^{n} \theta} \quad (10)$$

where $\hat{l}$ represent the GRNN estimated (x, y) coordinates of PD location associated with the observation fingerprint f, l_i are the training target (location coordinates coordinates) and n is the number of training samples.

The weights of the pattern neuron to D-summation neuron connections are set to unity, while weights of the pattern neurons to S-summation neuron connections are set to the output value of the training samples.

The smoothing parameter σ is the only free parameter in the GRNN network that needs to be determined. This is not trivial however. Too large a value of smoothing parameter can result in loss of detailed information in the estimated density. Too small a value of the smoothing parameter introduces disturbance caused by localized features or noise in the estimated density. The optimal value of the smoothing parameter (σ = 0.14) is obtained by v-fold cross-validation. The original training data is randomly divided into v segments (folds). In each case, one of the v subsets is used as validation data and the remaining are used for training. This process is then successively applied to all choices of v segments and the computed errors are averaged to yield a measure of the performance of the model. The above steps are repeated for various σ and the value achieving the lowest average error is selected as the optimal value. In this work, a 10-fold cross-validation is used.

6. Results and Discussions

In order to validate the proposed PD localisation technique, the proposed technique is analysed and compared with the well-known techniques often used for fingerprinting applications: KNN and WKNN technique. The independent dataset collected from 32 distinct locations as described in section 3.2 was used to evaluate the location determination performance of the models. The evaluation was performed as emulated localisation. This means the fingerprints (ratios of RSS) are presented to the trained/optimised models as inputs and the returned location (x, y) estimates are compare to ground truth locations. Accuracy, precision and computational complexity were used to analyse and compare the models developed.

6.1. Location Estimation Accuracy and Precision

Results of emulated localisation of the models are given in Table 1. Accuracy here is given in terms of root mean squared error (RMSE) of location error. In this result, it was observe that the proposed GRNN model provides the highest accuracy which can be attributed to the fact that GRNN always converges to global minimal. GRNN model achieves RMSE of 1.81 m, compare to 2.12 m for KNN and 2.06 m for WKNN models. ANN models (MLP and GRNN) are shown to be robust against noise with lower standard deviation (1.04 and 1.07) and maximum error (3.92 and 4.38) compare to KNN models.

Models	RMSE (m)	Standard deviation (m)	75th percentile (m)	Maximum error (m)
KNN	2.12	1.25	2.77	5.59
WKNN	2.06	1.29	2.95	4.67
MLP	2.07	1.04	2.86	3.92
GRNN	1.81	1.07	2.44	4.38

Table 1. Model location accuracy

To provide a more detailed analysis, error distributions of the models performance is given in Fig. 9. This result is used to explain the precision performance of the PD localisation models. In these results, MLP and GRNN provide the highest precision performance, achieving approximately 83% and 87% location precision within 3 m respectively. This implies that 87 out of 100 PD locations were estimated with error distance not greater than 3 m. KNN and WKNN models show lower location precisions of 78% and 77% within 3m. Moreover, when the 75% of location error is considered as shown in Table 1, GRNN has error not more than 2.44 m.

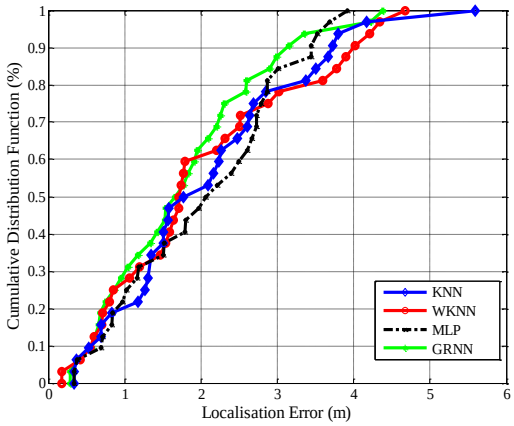


Fig. 9. *CDF of model localization errors*

The boxplot in Fig. 10 reveals the variation of localization error across the models. GRNN model shows a lower interquartile range with its localization errors skewed to the lower error bound, compared to KNN, WKNN and MLP models.

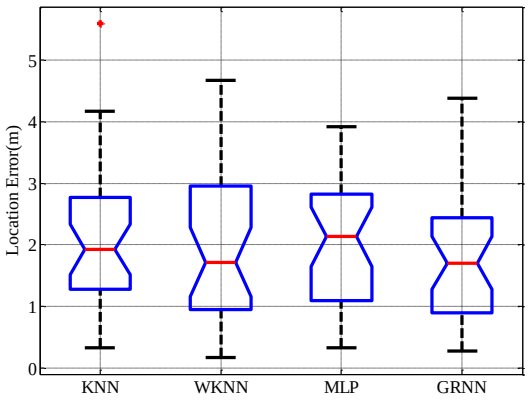


Fig. 10. *Boxplot of model localization errors.*

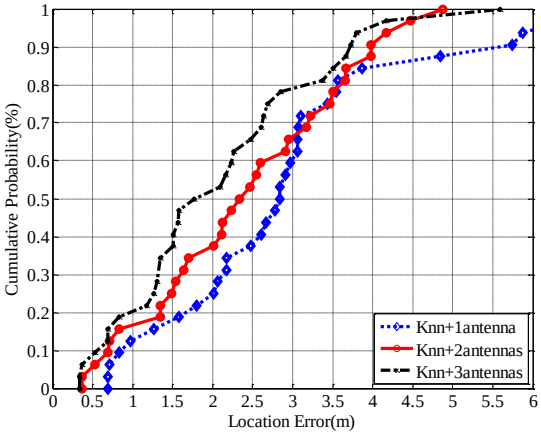
The application of KNN to PD location using radio fingerprinting limits the accuracy of the location estimate to the resolution of the grid points and the size of the area considered. It performs well for small areas, but for larger grids, the computational burden grows with size and the accuracy is often compromised.

6.2. Computation Complexity

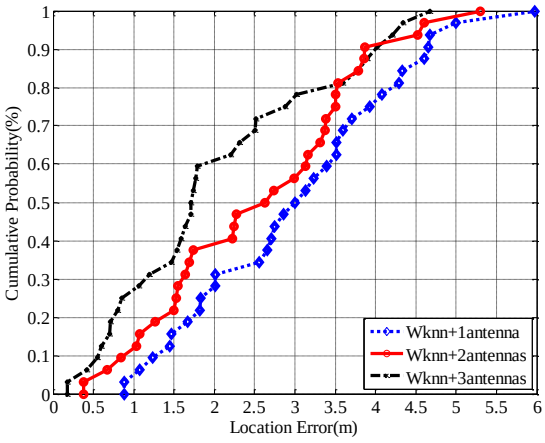
The computational complexity of the localisation models is assessed based on the dimension of the SSR vectors (P), the number of training or reference points (M) and the parameters of the algorithms used. KNN model creates a database of all training samples and in the localisation phase, it searches for nearest neighbours by comparing the distances between the test point and all training data. For every request of localisation, the computational complexity is therefore $O(KPM)$. Given the number of hidden layers (H), number of neurons for each hidden layer (F) of ANN and with the assumption that the evaluation of the activation function is negligible, the computational complexity for the generalisati on of artificial neural network, is $O(max\{H,P\}^2 \times F)$. It can be seen that the computational complexity of the KNN-based localisation model grows with increase in P, M and K. However, the complexity of our proposed model grows with only an increase in P since it has a fix structure. Therefore, our GRNN model is of low complexity compare to KNN models especially in cases where $M \gg P$.

6.3. Impact of Number of Antennas

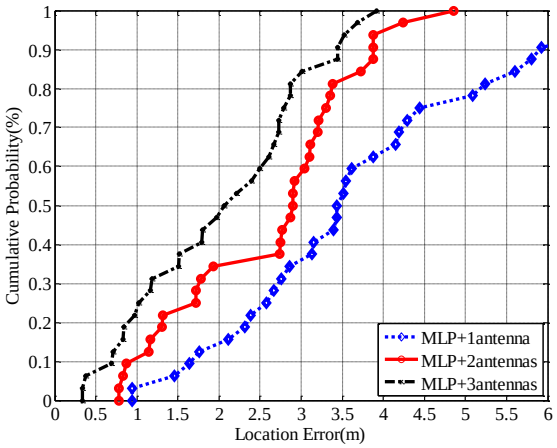
Radio fingerprinting technique relies on the unique RF signature created at each location in the propagation space. The number of possible unique signatures in turn depends on the number of receiving antennas. In order to evaluate the impact of number of antennas on the localisation accuracy, the data collected as describe in section 3.2 is used. First, SSR values on one antenna was used to infer PD locations using each model developed in this work. We then used SSR values on two antennas and finally SSR values on the three antennas. Fig. 11 (a-d) shows the impact of the number of receiving antennas on PD localization accuracy with each of the algorithms used in this paper. It was found that for each algorithm, there is a steady improvement in localization accuracy as the number of antennas increases from 1 through 3. In GRNN model the median and the 75th percentile localization error were reduced by 42 % and 34 % respectively when 3 antennas were used instead of 2. This indicates that in practical reality where a concentration of sensors would be deployed in electrical substation for PD localisation there would be a corresponding increase in localization accuracy.



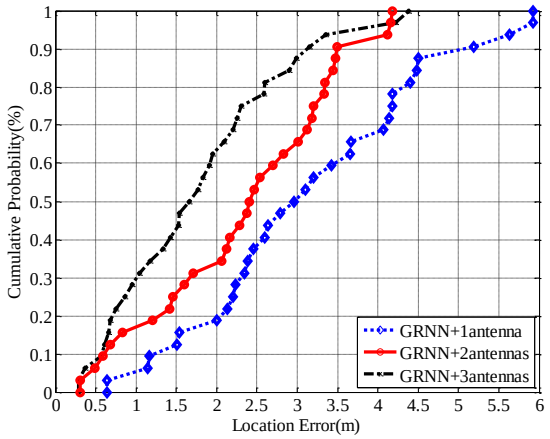
(a) *KNN*



(b) *WKNN*



(c) *MLP*



(d) *GRNN*

Fig. 11. *Location error with varying number of antennas*

7. Acknowledgments

The authors acknowledge the Engineering and Physical Sciences Research Council for their support of this work under grant EP/J015873/1 and Tertiary Education Trust Fund (TETFund) Nigeria.

8. Conclusion

A proof of concept WSS for condition-based monitoring in an electrical substation has been described. Specifically, we focus on the performance of machine learning algorithms to improve the performance of this distributed sensing application. The proposed algorithm is based on the use of PD signal strength ratio as location fingerprint rather than the absolute RSS. Two variants of the ANN learning algorithm (MLP and GRNN) are used to model the nonlinear relationship between the location fingerprints and the location coordinates of the PD source. The performance of the MLP and GRNN models have been evaluated and compared to the KNN models. Results show that MLP and GRNN models yield superior performance. The performance of the GRNN model (RMSE of 1.81 m) is of particular interest since it can function with a relatively small training set; and does not require computationally expensive training (iterative back propagation). As such particularly suitable for this application which uses computationally constrained hardware. Furthermore, it fits well with contemporary trends in Internet of Things (IoT) and Big Data towards edge processing. The algorithm's performance demonstrates that PD localization with adequate accuracy at reduced computational cost is indeed achievable. The results presented in this paper were obtained using only signal strength measurements recorded by 3 antennas. Future work will exploit and integrate other signal parameters for improved PD localization.

References

[1] Maddumage, S. K., Li, S., Pathirana, P., Williams, G.: "Entropy-based method to quantify limb length discrepancy using inertial sensors," *IET Wireless Sensor Systems*, vol. 8, no. 1, pp. 10 - 16, 2017.

[2] Di Franco, F., Tachtatzis, C., Atkinson, R. C., Tinnirello, I., Glover, I. A.: "Channel estimation and transmit power control in wireless body area networks," *IET Wireless Sensor Systems*, vol. 5, no. 1, pp. 11 - 19, 2014.

[3] Kwong, K. H., Wu, T. T., Goh, H. G., *et al.:* "Implementation of herd management systems with wireless sensor networks," *IET wireless sensor systems*, vol. 1, no. 2, pp. 55--65, 2011.

[4] Dai, X., Mitchell, J. E., Yang, Y., *et al.:* "Development and validation of a Simulator for Wireless Data Acquisition in Gas Turbine Engine Testing (WIDAGATE)," *IET Wireless Sensor Systems*, vol. 3, no. 3, p. 183 – 192, 2013.

[5] Saeed, B. I., Upton, D. W., Vieira, M. F. Q., *et al.:* "A Supervisory System for Partial Discharge Monitoring," *URSI Radio Science Conference*, Gran Canaria, Spain, 2018.

[6] Illias, H. A., Tunio, M. A., Bakar, A. H. A., Mokhlis, H., Chen, G.: "Partial discharge phenomena within an artificial void in cable insulation geometry: experimental validation and simulation," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 23, no. 1, pp. 451--459, 2016.

[7] Gao, W., Ding, D., Liu, W.: "Research on the typical partial discharge using the UHF detection method for GIS," *IEEE Transactions on Power Delivery*, vol. 26, no. 4, pp. 2621--2629, 2011.

[8] Zeng, F., Tang, J., Huang, L. Wang, W.: "A semi-definite relaxation approach for partial discharge source location in transformers," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 22, no. 2, pp. 1097--1103, 2015.

[9] Hou, H., Sheng, G., Jiang, X.: "Localization Algorithm for the PD Source in Substation Based on L-Shaped Antenna Array Signal Processing," *IEEE Transactions on Power Delivery*, vol. 30, no. 1, pp. 472-479, 2015.

[10] Portugues, I. E., Moore, P. J., Glover, I. A., *et al.:* "RF-based partial discharge early warning system for air-insulated substations," *IEEE Transactions on Power Delivery*, pp. 20--29, 2009.

[11] Moore, P. J. Portugues, I. E., and I. A. Glover, "Radiometric location of partial discharge sources on energized high-Voltage plant," *IEEE Transactions on Power Delivery*, vol. 20, no. 3, pp. 2264-2272, 2005.

[12] Iorkyase, E. T., Tachtatzis, C., Lazaridis, P., Glover, I. A., Atkinson, R. C.: "Radio location of partial discharge sources: A support vector regression approach," *IET Science, Measurement & Technology*, vol. 12, no. 2, pp. 230 - 236, 2018.

[13] Gui, L., Yang, M., Fang, P., Yang, S.: "RSS-Based Indoor Localization Using Multiplicative Distance-correction Factor," *IET Wireless Sensor Systems*, vol. 7, no. 4, pp. 98 - 104, 2017.

[14] Sinaga, H. H., Phung, B. T., Blackburn, T. R.: "Partial discharge localization in transformers using UHF detection method," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 19, no. 6, pp. 1891--1900, 2012.

[15] Ishimaru, H., Kawada, M.: " Locating multiple partial discharge sources using MAP estimation and ray tracing," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 9, no. s1, pp. s1-s7, 2014.

[16] Upton, D. W. Saeed, B. I. Mather, P. J. *et al.:* "Low power radiometric partial discharge sensor using composite transistor-reset integrator," *IEEE Transactions on Dielectrics and Electrical Insulation*, 2018.

[17] Upton, D. W., Haigh, R. P., Saeed, B. I., *et al.:* "Low power high-speed folding ADC based partial discharge sensor for wireless fault detection in substations," *URSI Radio Science Conference*, Gran Canaria, Spain, 2018.

[18] Alvarez, F., Ortego, J., Garnacho, F., Sanchez-Uran, M. A.: "A clustering technique for partial discharge and noise

sources identification in power cables by means of waveform parameters," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 23, no. 1, pp. 469--481, 2016.

[19] Genming, D., Zhang, J., Zhang, L., Tan, Z.: "Overview of received signal strength based fingerprinting localization in indoor wireless LAN environments," *IEEE 5th Int. Symp. on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications (MAPE)*, Chengdu, China, Oct. 2013, pp. 160-164.

[20] Kaemarungsi, K., Krishnamurthy, P.: "Modeling of indoor positioning systems based on location fingerprinting," *Twenty-third AnnualJoint Conf. of the IEEE Computer and Communications Societies INFOCOM*, Hong Kong, China, March 2004, pp. 1012-1022.

[21] Patwari, N., Hero, A. O., Perkins, M., Correal, N. S., O'dea, R. J.: "Relative location estimation in wireless sensor networks," *IEEE Transactions on signal processing*, vol. 51, no. 8, pp. 2137 - 2148, 2003.

[22] Kushki, A., Plataniotis, K. N., Venetsanopoulos, A. N.: "Kernel-based positioning in wireless local area networks," *IEEE transactions on mobile computing*, vol. 6, no. 6, pp. 689 - 705, 2007.

[23] Burba, F., Ferraty, F., Vieu, P.: "k-Nearest Neighbour method in functional nonparametric regression," *Journal of Nonparametric Statistics*, vol. 21, no. 2, pp. 453--469, 2009.

[24] Kuriakose, J., Naveenbabu, J. V., Shahid, M., Shetty, A.: "Analysis of Maximum Likelihood and Mahalanobis Distance for Identifying Cheating Anchor Nodes," *Inter. Conf. on Emerging Research in Computing, Information, Communication and Applications*, Dec. 2014.

[25] Hamza, L., Nerguizian, C.: "Neural network and fingerprinting-based localization in dynamic channels," *IEEE Inter. Symp. on Intelligent Signal Processing, WISP*, Budapest, Hungary, Aug. 2009, pp. 253-258.

[26] Werle, P., Akbari, A., Borsi, H., Gockenbach, E.: "Partial discharge localisation on power transformers using neural networks combined with sectional winding transfer functions as knowledge base," *Inter. Symp. on Electrical Insulating Materials*, Himeji, Japan, Nov. 2001, pp. 579-582.

[27] Yang, J., Zeng, X., Zhong, S., Wu, S.: "Effective neural network ensemble approach for improving generalization performance," *IEEE transactions on neural networks and learning systems*, vol. 24, no. 6, pp. 878--887, 2013.

[28] Specht, D. F.: "A general regression neural network," *IEEE Transactions on Neural Networks*, vol. 2, no. 6, pp. 568--576, 1991.

[29] Chun-Yao, L., Yan-Lou, H.: "Wind prediction based on general regression neural network," *Second Inter. Conf. on Intelligent System Design and Engineering Application (ISDEA)*, Hainan, China, Jan. 2012, pp. 617-620.