Investigation of RBF kernelized ANFIS for Fault Diagnosis in Wireless Sensor Networks

Rakesh Ranjan Swain[1], Tirtharaj Dash[2], and Pabitra Mohan Khilar[1]

[1] Department of Computer Science and Engineering
National Institute of Technology Rourkela, India
[2] Department of Computer Science and Information Systems
Birla Institute of Technology and Science Pilani, Goa Campus, India

Abstract. Wireless sensor networks (WSN) are often inaccessible to human and are at least deployed in such environment such as deep forest, various hazardous industries, hilltop, and sometimes underwater. The occurrence of failures in sensor networks is inevitable due to continuous or instant change in environmental parameters. A failure may lead to faulty readings which in turn may cause economic and physical damages to the environment. In this work, a thorough investigation has been conducted on the application of adaptive neuro-fuzzy inference system (ANFIS) for automated fault diagnosis in WSN. Further, a kernelized version of ANFIS has also been studied for the discussed problem. To avoid the model's undesired biases towards a specific type of failure, oversampling has been done for multiple version of the ANFIS model. This study would serve as a guideline for the community towards the application of fuzzy inference approaches for fault diagnosis in sensor networks. However, the work focuses on the automated fault diagnosis in open air WSN and has no applicability in underwater sensor network systems.

Keywords: Fault diagnosis, Wireless Sensor Networks, Kernelization, ANFIS

1 Introduction

Wireless sensor network (WSN) is a set of sensor nodes that are often low in cost, power, memory storage, and are small sized and sometimes multi-functional. A sensor node contains sensory unit, processor, memory storage, actuator, and power supply unit. The sensor nodes are sometimes equipped with multiple sensory units to collect various environmental parameters for further processing. It is well known that WSN has found their applicability in numerous applications such as target tracking, environmental monitoring, health monitoring [1]. Generally, sensor nodes are densely deployed in an unstructured manner for the purpose of sensing, processing and communicating with each other. In an unstructured deployment in a comparatively human inaccessible environment, failure of sensor nodes is inevitable. A failure in a WSN leads to faulty readings which in turn may cause economic and physical damages to the environment.

For example, failures that might occur specifically in human health monitoring environment or in chemical power plants are quite prominent. Diagnosis of faults in WSN is quite a well-studied research for the community [2]. There are four different types of sensor faults that can be seen in WSN such as hard fault, soft-permanent fault, intermittent fault, and transient fault. The latter three types of faults can be grouped in the category of soft fault.

Accurate diagnosis of fault in a real-time environment from the sensor readings data is quite a challenge for the community. To the best of our belief, the majority of fault diagnosis works focused on the detection of two broad types of faults such as (a) hard fault, (b) soft fault. These works did not focus more on the three sub-categories of soft faults as mentioned earlier. The present work focuses on the automated diagnosis of the soft faults. Machine learning (ML) based techniques such as Neural Networks are very popular approaches in the field of automated fault diagnosis, for example, a work presented in [3,4], and handling uncertainty by incorporating fuzzy logic[5]. Further, the works carried out in this field have been identified in the section 5.

In the present work, we investigate the application of adaptive neuro-fuzzy inference system (ANFIS) for automated fault diagnosis in WSN. Our proposed method has been investigated for all three different types of soft faults. Further, we equip the ANFIS module with radial basis kernel (RBF) to study whether such an approach could improve the performance of conventional ANFIS technique. In real-time WSN, the occurrence of few categories of soft faults is minimal that offers another challenge for an ML approach to accurately diagnose the faults from the available sensor readings. In this work, we also study the automatic synthesis of faulty sensor readings from the available data using a relative distance based oversampling technique such as SMOTE.

The rest of the paper is presented as followed. We present our proposed kernelized ANFIS approach in section 2. The experimental design has been studied in the section 3. Results obtained by the proposed approach have been discussed in the section 4. The work has been concluded in the section 6.

2 Kernelized ANFIS

ANFIS is a special class of multi-layer adaptive networks that incorporate both neural networks (NN) and fuzzy logic principles to model uncertain systems [6] such as WSN. Neural networks are supervised learning models that approximate the underlying transformation function from the historical input-output data pairs. The function suggested by the NN could further be used for future predictions where the output is unknown. In fuzzy logic, a rule-base drawn from the available historical data and the control signal is generated by firing the rule base. ANFIS makes the selection of the rule base more adaptive to the situation and the problem under consideration.

Fig. 1 shows the architecture of a typical ANFIS with two inputs, four rules and one output for the first-order Sugeno fuzzy model, where each input is assumed to have two associated membership functions (MFs). For a first-order

Sugeno fuzzy model, a typical rule can be written as

$$\text{If } x \text{ is } A_i \text{ and } y \text{ is } B_j, \text{ then } f_{ij} = a_{ij}x + b_{ij} + c_{ij} \tag{1}$$

where, A_i and B_j are the membership functions (MF) for the inputs x and y respectively. The constants a_{ij}, b_{ij} and c_{ij} are consequent parameters [6]. A set of such rules can be obtained by changing the inputs or the MF. The functionality of each layer can be briefed as follows[3]. The layer 1 is an adaptive layer that generates the membership grades for the inputs using the MF. The MF parameters in this layer are called the premise or antecedent parameters. The layer 2 computes the firing strength of each of the rules by doing simple multiplication or minimum operation, usually denoted as Π. Layer 3 computes the normalized firing strength by taking individual firing strengths from the layer 2. Each node in the layer 4 an adaptive node and outputs the product of the normalized firing strength and the first order polynomial for the first order Sugeno model. The final layer (layer 5) computes the output by taking the summation of the outputs from the previous layer. Learning of the parameters in layer 1 (premise parameters) and layer 4 (consequent parameters) happens by the squared error signal generated at the output.

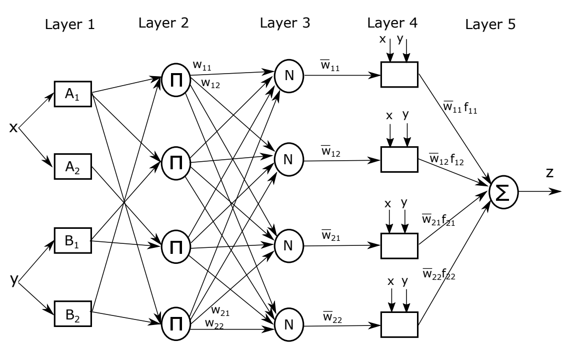


Fig. 1: ANFIS architecture

The hypothesis that we are testing in this work is that a problem, when cast in higher dimensional feature space is more likely to be separable than that in a lower dimensional space [7]. The present work investigates this hypothesis by adding a nonlinear transformation layer where the original inputs are transformed to a new feature space in a higher dimension. Each computational unit j in this layer does a nonlinear transformation (f_t in Fig.2) using RBF kernel defined in Equation (2) of the input data pattern which is denoted as

[3] For page constraints, we have intentionally omitted major mathematical background behind the functionality of each layer. Readers are advised to go through the main ANFIS article by Professor Jang [6].

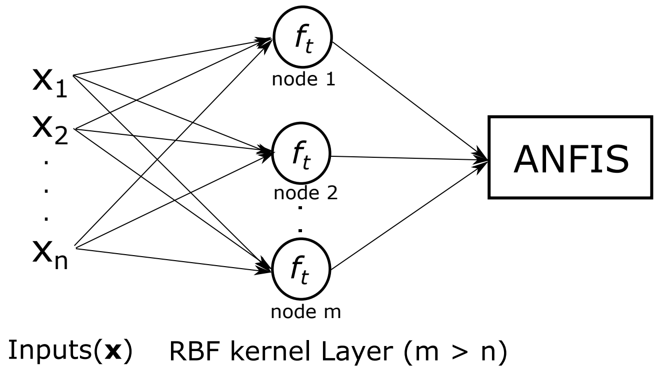


Fig. 2: Kernelized ANFIS architecture (The ANFIS block this figure uses the layer 1 through to layer 5 of Fig. 1 where the inputs to the ANFIS are kernelized using Radial basis functions)

$\mathbf{x} = (x_1, x_2, \ldots, x_n)$.

$$f_t = \varphi(\mathbf{x}) = \varphi(\mathbf{x} - \mathbf{x}_j) = \exp\left(-\frac{1}{2\sigma_j^2}||\mathbf{x} - \mathbf{x}_j||^2\right), j = 1, 2, \ldots, m \qquad (2)$$

In the Equation (2), $\mathbf{x}_j$ is the center of the data points; $||\cdot||$ represents the Euclidean distance. The number of neurons in the RBF transformation layer, m has been decided by applying the k-means clustering [8] on the available training data with only the input features. It is believed that patterns sharing common characteristics would be grouped together. If there are N patterns in a dataset, we decided that the value of k in k-means clustering is strictly lesser than N ($k < N$), and strictly higher than the number of features of the data, that is $k > n$ to support our initial hypothesis of transformation to higher dimension. It should be noted that $m = k$.

In reality, sensor data from the faulty sensors are available in few chunks than the normal data. These sensors might belong to any of the three classes of the soft faults. We employed synthetic minority oversampling technique (SMOTE) [9] to generate the sensor data that are similar to the available data. This approach also adds an investigation of the synthetic sampling technique to the fault diagnosis in WSN. The final algorithm has been presented in the following.

3 Design of Experiment

The present work is a simulation-based experiment where the data is obtained from real-world experimental setup as explained further.

The experiment was conducted in an outdoor environment. In this experiment, 10 sensor nodes were deployed in an area of 5×5 m^2. Fig. 3 shows the deployment of all these sensor nodes. Each sensor module is designed by Arduino

Data: Sensored data from the central receiver
Result: Fault diagnosis results on unseen test data
Step 0: Initialization and setting necessary ANFIS parameters and m;
Step 1: Partition the available data into training data (D_{train}) and testing data (D_{test}) by random sampling of the patterns from the whole dataset;
Step 2: Obtain the oversampled data, $D_{osTrain}$ = SMOTE(D_{train});
Step 3: Apply k-means algorithm on $D_{osTrain}$ without the class labels and obtain centroids of each clusters;
Step 4: Apply RBF kernelization scheme to get the data ($D_{rbfOsTrain}$) for ANFIS training;
Step 5: Train ANFIS on $D_{rbfOsTrain}$ using conjugate gradient method;
Step 6: Test the trained ANFIS using the (D_{test});

Algorithm 1: RBF kernelized ANFIS with oversampling for fault diagnosis in WSN

Uno embedded and At mega 8-bit micro-controller, DHT11 temperature sensor, MRF24J40MA transceiver, and 2200 mAh battery. All these sensor nodes are communicated to the base station with a gateway node. The diagnosis program is run in the base station. All the sensor nodes and the gateway node are presented in the transmission range of each other. The transmission power is set –16 dBm for the communication range in between $5m$ to $6m$. The gateway node is presented in between $6m$ distance.

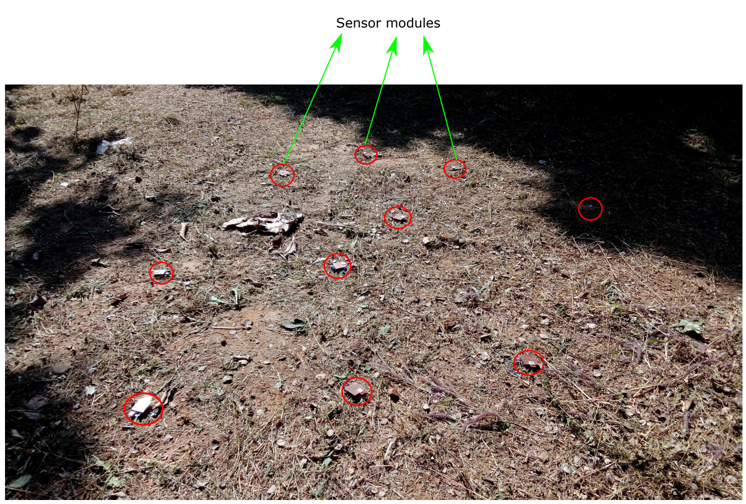


Fig. 3: Experimental setup in an outdoor environment. A gateway node communicates with these sensor modules.

Initially, all the nodes were considered fault free in nature. The sensor nodes temperature data were collected every hour of a daytime. Then after analyzed the normal sensor data, we set the minimum threshold temperature value θ_1, and maximum threshold value θ_2. The sensor nodes data within the θ_1 and θ_2 is considered as fault free and exceeding the range, it is considered as faulty. Let x_i be the sensor data from sensor node n_i at t^{th} time instance. So, the sensor data is denoted as $x_i(t)$. For sensor node n_i, we collects the sensor data $\mathbf{x} = (x_i(t)); t = 1, \ldots, T$. From the time $t = 1$ to T, all the sensor data were analyzed. The data $x_i(t)$ follows the normal distribution $\mathcal{N}(\mathbf{x}_i; \sigma_i^2)$, where $\mathbf{x}_i$ is the actual data and σ_i^2 is the variance of fault data at n_i sensor node (noise to make the node faulty). So this value, we are represented as in the Eq. (3). Similarly, depending on the probability of noisy data, the faulty sensor readings could be generated for different types of faults.

$$\mathbf{x}_i(t) = \mathbf{x}_i + e_i(t);\ e_i \sim \mathcal{N}(\mathbf{x}_i; \sigma_i^2) \tag{3}$$

All the ANFIS based simulations have been conducted in MATLAB R2017a on a quad-core system with 4GB RAM. After obtaining the sensor readings (including the faulty sensor readings data), the whole dataset was partitioned for the training and testing simulation of various ANFIS approaches. The train and test partitioning were done by randomly selecting the patterns as per 70%: 30% ratio respectively. For all the ANFIS experiments, the number of training epochs was set to 100. The dataset summary has been reported in the Table 1.

Table 1: Available sensor data summary (N: Total number of patterns, $N(class = i)$: Number of patterns of ith class)

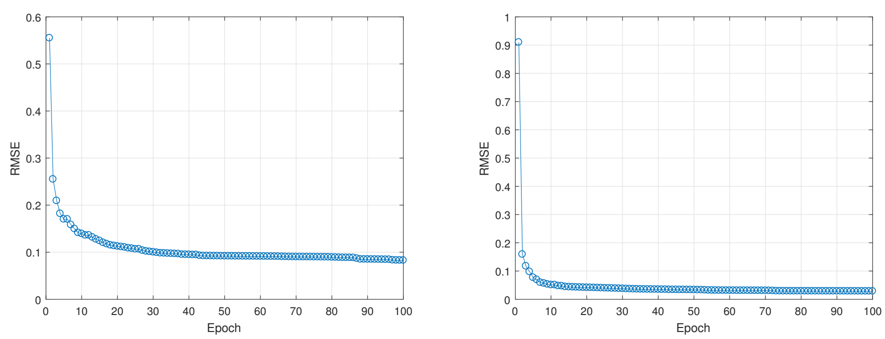
Data	N	$N(class = 1)$	$N(class = 2)$	$N(class = 3)$	$N(class = 4)$
Original	5000	2000	500	1500	1000
Train (D$_{train}$)	3500	1398	333	1064	705
Test (D$_{test}$)	1500	602	152	452	294

It should be noted that the oversampling was only applied to the training dataset. The test dataset remains intact throughout the experiment except for the RBF kernelization experiment in which both the training and testing are to be transformed to a new higher dimensional feature space where the centroids were obtained using the k-means clustering on the training data only. Further, the results are analyzed in the following section.

4 Results

The spread of the RBF kernel, σ plays an important role as an amplifier of the distance between a pattern and the center of the kernel. If this distance is much larger than σ, the kernel function tends to be very minimal. Moreover, it has a high impact on the decision rules–specifically, boundary–that can affect the prediction of the testing data. Therefore, the results for two different values of the spread is noted (a) $\sigma = 5$, (b) $\sigma = 10$. The number of sub-grouping(clustering) in the available training data, m has been varied from the 15 to 25 with equal separation of 5 to investigate the effect of the transformed feature space in a slightly higher dimension ($m \approx n; m = 15, n = 10$) and a dimension as large as twice the input feature space ($m > 2n; m = 25, n = 10$).

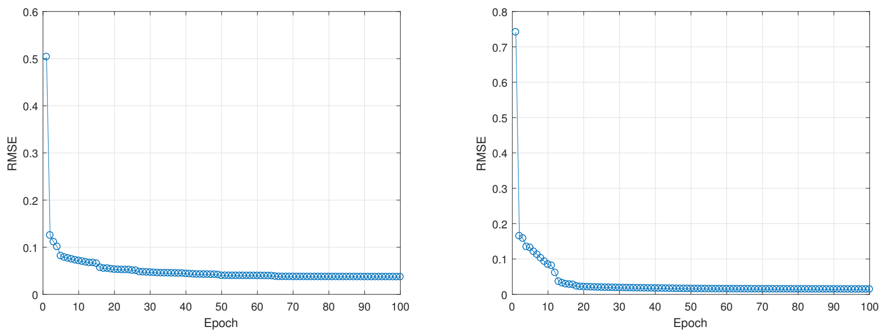
We used root-mean-squared error (RMSE) as a performance evaluation measure of ANFIS models. A few sample training error convergence curve are presented in the Figures 4 through to 6. For the further investigate the quality of fault diagnosis in terms of the faulty nodes prediction rate, we used the accuracy measure. We also report the prediction rate after training and after testing to clearly investigate the specialization and generalization capability of the developed ANFIS models. These results are summarized in Table 2 where m is the size of the transformed feature space and σ is the spread of the RBF kernel.



(a) ANFIS (b) ANFIS-SMOTE

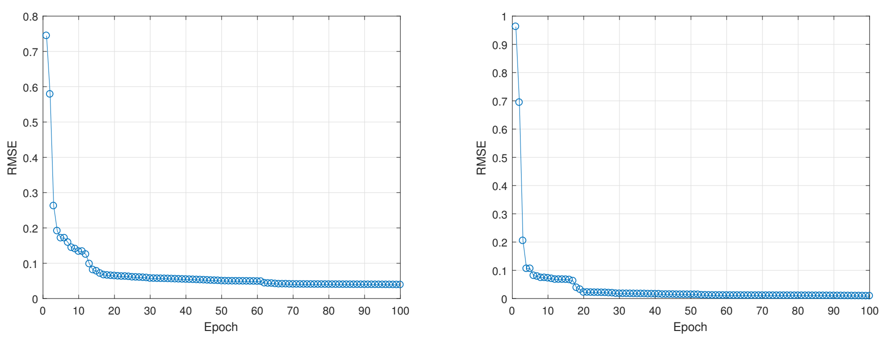
Fig. 4: Error convergence during training

The present investigation suggests that a higher dimensional transformation of the input feature space prior to the learning rules by the ANFIS might not lead to an improved predictive capability of the models for the present problem on the faulty diagnosis from the available data. It can be revealed from the data reported in the Table 2. The conventional ANFIS and the ANFIS with SMOTE oversampling are quite comparable with each other based on their presently achieved generalization capability on the test dataset. However, the RBF kernelized ANFIS with and without oversampling are very highly specialized towards the training data that is leading to over-fitting and in turn poor generalization. Hence, such a model might not be preferable for the present problem. However, further experiments must be conducted carefully before concluding in a general sense.



(a) ANFIS ($m = 15, \sigma = 5$) (b) ANFIS ($m = 15, \sigma = 5$)-SMOTE

Fig. 5: Error convergence during training



(a) ANFIS ($m = 25, \sigma = 10$) (b) ANFIS ($m = 25, \sigma = 10$)-SMOTE

Fig. 6: Error convergence during training

5 Related works

Several different soft computing or statistical techniques have been employed recently for studying the fault diagnosis in WSN. A statistical approach called neighboring coordination method was used where the authors used the statistical variance of the sensor readings for predicting the fault [10]. A Takagi-Sugeno-Kang (TSK) based fuzzy inference system was applied for detection of soft permanent faults and transient faults in WSN [11]. Fuzzy transformation based multi-layer feed-forward neural network [12] has been successfully implemented for automated fault diagnosis in WSN [13]. However, this work used a single membership function to generate a single transformed feature instead of three different features as in our present work. There are many different approaches to automated fault diagnosis in WSN that use different kinds of neural net-

Table 2: Fault diagnosis (accuracy of prediction) scores achieved by various ANFIS approaches (RMSE has been noted for Epoch1 and the Epoch 100 shown as transition '$\rightarrow$')

Method	Training (%)	Testing (%)	RMSE
ANFIS	95.88	94.80	$0.5549 \rightarrow 0.0827$
ANFIS-SMOTE	98.55	94.20	$0.9098 \rightarrow 0.0285$
ANFIS ($m = 15, \sigma = 5$)	98.17	16.46	$0.6201 \rightarrow 0.0369$
ANFIS ($m = 15, \sigma = 10$)	93.60	16.80	$0.7435 \rightarrow 0.1135$
ANFIS ($m = 20, \sigma = 5$)	79.82	34.53	$0.5874 \rightarrow 0.4002$
ANFIS ($m = 20, \sigma = 10$)	88.51	20.20	$0.7795 \rightarrow 0.2025$
ANFIS ($m = 25, \sigma = 5$)	99.14	32.60	$0.5292 \rightarrow 0.0179$
ANFIS ($m = 25, \sigma = 10$)	97.88	29.86	$0.7442 \rightarrow 0.0381$
ANFIS ($m = 15, \sigma = 5$)-SMOTE	99.28	17.93	$0.7415 \rightarrow 0.0141$
ANFIS ($m = 15, \sigma = 10$)-SMOTE	99.47	18.53	$0.9703 \rightarrow 0.0099$
ANFIS ($m = 20, \sigma = 5$)-SMOTE	99.60	38.73	$0.9095 \rightarrow 0.0076$
ANFIS ($m = 20, \sigma = 10$)-SMOTE	99.33	22.26	$1.0170 \rightarrow 0.0111$
ANFIS ($m = 25, \sigma = 5$)-SMOTE	99.85	35.00	$0.7919 \rightarrow 0.0029$
ANFIS ($m = 25, \sigma = 10$)-SMOTE	99.51	30.26	$0.9625 \rightarrow 0.0089$

works such as probabilistic neural network [14] and statistical techniques such as hypothesis testing [10]. Many different kinds of neural networks have been developed recently for use in fault diagnosis and could be referred to [15,16].

6 Conclusion

Many different research works present the success stories behind the application of various methods to automated fault diagnosis in WSN. The present work focused on a thorough investigation of the possible modification to fuzzy inference based investigation scheme using kernelized feature space. The achieved results suggested that the application of such a scheme might not be suitable for the problem under consideration because of the fact that it makes the inference engine highly specialized towards the training distribution and makes the machine weak in generalization which is the most important aspect of learning machines. On the basis of the obtained results, one could further study in the following possible directions.

Our primary results show that ANFIS with SMOTE without the RBF kernelization outperforms ANFIS during training. However, it has identical results in testing. This leaves us with a question to investigate while using other learning machines. Specifically, we should be interested in using oversampling for rule-generators such as decision tree and random forest classifiers. It could be hypothesized that an oversampling engine would generate missing information for the sensor dataset and could make it complete. Then a decision tree can be used to learn the set of rules for the input-output transformation. Further, such a model is not computationally expensive unlike sophisticated models like neural networks.

The results achieved due to the kernel trick in the ANFIS (and ANFIS with SMOTE) is worse than its non-kernelized counterpart because of poor generalization of the underlying learning machine. A straightforward question to investigate in future is to study the use of support vector machines (SVM) with kernel trick for automated classification of sensor faults. As discussed in the aforementioned point, such a method would also be computationally inexpensive as compared with various neural networks.

Furthermore, failures in sensor networks might occur in both sensor nodes and the links. It might be arguably correct to say that failures in linked sensor nodes would cause link failure. But, for instance, the primary focus is that such failures lead to the formation of islands or cuts in sensor networks. Earlier studies have successfully used simple graph theoretic approaches such as [17] for detection of cuts in WSN. Sometimes third party intrusion might cause such failures in the network. So, in such circumstances, a robust model must be designed for automated diagnosis of failures that would not only detect failures but also decide whether the failure was caused due to environmental damages or third party intrusion [18,19,20]. To the best of our knowledge, such a model that would deal with both these kinds of decisions has not been built.

References

1. Arampatzis, T., Lygeros, J., Manesis, S.: A survey of applications of wireless sensors and wireless sensor networks. In: Intelligent Control, 2005. Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation. pp. 719–724. IEEE (2005)
2. Mahapatro, A., Khilar, P.M.: Fault diagnosis in wireless sensor networks: A survey. IEEE Communications Surveys & Tutorials 15(4), 2000–2026 (2013)
3. Hou, L., Bergmann, N.W.: Novel industrial wireless sensor networks for machine condition monitoring and fault diagnosis. IEEE Transactions on Instrumentation and Measurement 61(10), 2787–2798 (2012)
4. Swain, R.R., Khilar, P.M.: Composite fault diagnosis in wireless sensor networks using neural networks. Wireless Personal Communications 95(3), 2507–2548 (2017)
5. Reddy, P.N., Dambekodi, S.N., Dash, T.: Towards continuous monitoring of environment under uncertainty: A fuzzy granular decision tree approach. CEUR Workshop Proceedings 1819 (2017)
6. Jang, J.S.: Anfis: adaptive-network-based fuzzy inference system. IEEE transactions on systems, man, and cybernetics 23(3), 665–685 (1993)
7. Cover, T.M.: Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. IEEE transactions on electronic computers (3), 326–334 (1965)
8. Hartigan, J.A., Wong, M.A.: Algorithm as 136: A k-means clustering algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics) 28(1), 100–108 (1979)

9. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: Smote: synthetic minority over-sampling technique. Journal of artificial intelligence research 16, 321–357 (2002)
10. Panda, M., Khilar, P.M.: Distributed self fault diagnosis algorithm for large scale wireless sensor networks using modified three sigma edit test. Ad Hoc Networks 25, 170–184 (2015)
11. Khan, S.A., Daachi, B., Djouani, K.: Application of fuzzy inference systems to detection of faults in wireless sensor networks. Neurocomputing 94, 111–120 (2012)
12. Dash, T., Behera, H.S.: A fuzzy mlp approach for non-linear pattern classification. arXiv preprint arXiv:1601.03481 (2015)
13. Swain, R.R., Khilar, P.M.: A fuzzy mlp approach for fault diagnosis in wireless sensor networks. In: Region 10 Conference (TENCON), 2016 IEEE. pp. 3183–3188. IEEE (2016)
14. Swain, R.R., Khilar, P.M., Bhoi, S.K.: Heterogeneous fault diagnosis for wireless sensor networks. Ad Hoc Networks 69, 15–37 (2018)
15. Ji, Z., Bing-shu, W., Yong-guang, M., Rong-hua, Z., Jian, D.: Fault diagnosis of sensor network using information fusion defined on different reference sets. In: Radar, 2006. CIE'06. International Conference on. pp. 1–5. IEEE (2006)
16. Moustapha, A.I., Selmic, R.R.: Wireless sensor network modeling using modified recurrent neural networks: Application to fault detection. IEEE Transactions on Instrumentation and Measurement 57(5), 981–988 (2008)
17. Swain, R.R., Dash, T., Khilar, P.M.: An effective graph-theoretic approach towards simultaneous detection of fault (s) and cut (s) in wireless sensor networks. International Journal of Communication Systems (2017)
18. da Silva, A.P.R., Martins, M.H., Rocha, B.P., Loureiro, A.A., Ruiz, L.B., Wong, H.C.: Decentralized intrusion detection in wireless sensor networks. In: Proceedings of the 1st ACM international workshop on Quality of service & security in wireless and mobile networks. pp. 16–23. ACM (2005)
19. Roman, R., Zhou, J., Lopez, J.: Applying intrusion detection systems to wireless sensor networks. In: IEEE Consumer Communications & Networking Conference (CCNC 2006) (2006)
20. Dash, T.: A study on intrusion detection using neural networks trained with evolutionary algorithms. Soft Computing 21(10), 2687–2700 (2017)