This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

A Smart Optimization of Fault Diagnosis in Electrical Grid Using Distributed Software-Defined IoT System

Ammar K. Al Mhdawi, *Member, IEEE*, and Hamed Saffa Al-Raweshidy, *Senior Member, IEEE*

Abstract*—Electrical power demands have increased significantly over the last years due to rapid increase in air conditioning units and home appliances per domestic unit, particularly in Iraq. Having an uninterrupted power supply is essential for the continuity of power-generated home services and industrial platforms. Currently, in Iraq, electrical power interruption has become a big concern to the utility suppliers. Despite successive attempts to put an end to this dilemma, the issue still prevails. One of the main factors in power outages in local zones is persistent faults in distribution transformers (DTs). DT is considered one of the main elements in the electrical network that is essential for the reliability of the grid supply. Due to the internal lack of monitoring system and periodic maintenance, DT is relentlessly subject to faults due to high overhead utilization. Therefore, in order to enhance the grid reliability, transformer health check, and maintenance practices, we propose a remote condition Internet of Things monitoring and fault prediction system that is based on a customized software-defined networking (SDN) technology. This approach is a transition to smart grid implementation by fusing the power grid with efficient and real-time wireless communication architecture. The SDN implementation is considered in two phases: one is a controller installed per local zone and the other is the main controller that is installed between zones and connected to the core network. The core network consists of redundant links to recover from any future fails. Furthermore, we propose a prediction system based on an artificial neural network algorithm, called distribution transformer fault prediction, that is installed in the management plane for periodic prediction based on real-time sensor traffic to our proposed cloud. Moreover, we propose a communication protocol in the local zone called local SDN-sense. The SDN-sense ensures a reliable communication and local node selection to relay DT sensor data to the main controller. Our proposed architecture showcases an efficient approach to handle future interruption and faults in power grid using cost-effective and reliable infrastructure that can predict and provide real-time health monitoring indices for the Iraqi grid network with minimal power interruptions. After extensive testing, the prediction accuracy was about 96.1%.

***Index Terms*—Fault prediction, Long Range-Internet of Things (LoRa-IoT), monitoring network, neural networks (NNs), sensors, smart grid, software-defined networks (SDNs).**

I. Introduction

GENERALLY, high power has to be generated and supplied to the domestic and industrial units on a 24/7 basis. The power source and distribution network of the electrical system have to be maintained continuously to provide nonstop electricity consumption. Traditional power grid relies on human operators to manage and monitor the status and efficiency of the grid, and coordinate supply and demand to ensure reliable stability of the power grid [1]. The significant increasing requirements for quality power management are implemented via deploying monitoring and control strategies all over the grid system. Traditional distribution transformers (DTs) have an average life of 20–25 years; however, most of these transformers are at the end cycle of their life and are posing an intermittent risk to the power grid system. The current monitoring system of the power grid in Iraq is only associated with major electrical parameters that provide no health check status on the internal components of the local distribution network. Lack of periodic maintenance and follow up checks is a major factor in these repetitive DT failures that is due to nonestablished visibility system. Therefore, a robust monitoring and prediction system is needed to establish real-time monitoring of each distribution unit of the local grid [2] by using software-defined networking (SDN) principle. SDN is a new programmable network concept paradigm that has been proposed recently to facilitate management and data steering of the network. SDN is the concept of separating the control plane from the data plane in which the forwarding hardware is segregated from the decision-making platform such as routing and control software [3]. The separation of the planes provides a flexible, programmable, and cost-effective network infrastructure. In the SDN network, the policies will be running on the controller only instead of running them on each device as in the traditional network. The controller will have a full overview of the network topology and all nodes can be configured from a single point of management. This approach will provide a robust management of large-scale network with less overhead. Each engine has a table called forwarding table that forwards on the basis of matching the incoming packet to the table. The communication between the SDN and OpenFlow switches is governed by the OpenFlow protocol. The OpenFlow protocol is a set of messages that are exchanged between the controller and the switches over a secure connected channel. The controller sends modification messages to the switch node such as add, modify, or remove entries from the forwarding table. When an incoming packet enters the OpenFlow switch, it maps the packet info to the forwarding table. If there is a match, then it forwards the packet to the designated port; otherwise, it sends a query request to the controller to request advice from the controller on where to send the packet. The SDN controller then consults its topology table and decide whether to send new rules or notify

Manuscript received February 11, 2019; revised April 26, 2019 and June 3, 2019; accepted June 4, 2019. *(Corresponding author: Ammar K. Al Mhdawi.)*
The authors are with Brunel University London, Uxbridge UB8 3PH, U.K. (e-mail: eepgaka1@brunel.ac.uk; hamed.Al-raweshidy@brunel.ac.uk).
Digital Object Identifier 10.1109/JSYST.2019.2921867

the switch to drop the packet. Furthermore, SDN has two main interfaces: one is northbound interface that is used to push configuration, read, install rules, and implement modifications on excising rules; and the second interface is the southbound interface that is used by the controller to push rules and modification to the lower level nodes or OpenFlow switches [4]–[8].

To implement a structural health monitoring system for the DTs, a wireless sensor network (WSN) is considered. A WSN is a network that is constructed using a large number of distributed nodes, where each node consists of a specific sensor that detects a physical condition of an object such as temperature, heat, liquid levels, and pressure. Sensor nodes monitor the condition and transmit the data along to other nodes until it reaches the management node or gateway that represents the collection of all data. The sensors are powered by either a fixed power source such as batteries or by using an energy harvesting technology such as solar, thermal, or kinetic [9]. However, wireless sensors are limited by over-the-air transmission obstacles that could hinder the transmission rate of data comparable to the wired network systems. Installation of sensors on electrical grid components can provide an immediate status of components condition, which helps in understanding how the grid can handle a certain electrical load and can provide early fault alert with minimum low cost of repair. The result of using WSN correlate with the increase in profitability and stability of the electrical grid.

Traditional tools are not always capable of achieving efficient accuracy and reliability regarding fault classification of DTs. The process of identifying faults in the DT components is significantly crucial for the continuity of the power supply. It can help in reducing the number of unexpected faults, reduce maintenance cost, and help in extending the life cycle of the transformer [10]. Additionally, by using a smart intelligent system, it becomes a coherent process to assist in analysis and fault classification of the operational transformer based on its current load status. Moreover, neural networks (NNs) have been greatly used in the electrical power network for predicting power production and estimating power demands.

Recently, researchers have been using statistical modeling and methods to evaluate and analyze the behavior of the power grid network. However, NN is considered a new approach in prediction compared to conventional prediction methods. The strength of NNs is that they do not need any assumptions and they use previous historical data to generate prediction by optimizing the nonlinearity issues in the system. The prediction is done by constructing a complex relationship between the input and the output by applying rounds of training and optimization on a given dataset [11]. Moreover, NNs consist of neurons or perceptrons that are interconnected with each other via links. There are three main layers in a NN: the input layer; hidden layer; and output layer. A perceptron has multiple inputs to it with weights for each link. Details of the proposed NNs architecture are described later in Section III.

II. Related Work

SDNs [12] have played a significant role in reconstructing the network architecture to less complex and flexible elements in terms of deployment and flexibility. Moreover, NN has been establishing a solid ground in many sectors by predicting the status of system behavior and provide accurate predictions based on historical data. Nonetheless, researchers have been working on different modeling techniques to implement the NN in power grid to predict power supply performance and fault diagnosis. In the following, we list some of the work that was implemented by researchers to put solutions for some of the challenges and concerns that occurred in the power grid.

Grid component faults are significant problems in power distribution. For that, Zhang *et al.* [13] proposed a method for prediction of the trip fault using long short-term memory and support vector machines, which are a high margin classifier in NNs. The data were captured with the long short term memory (LSTM) network with a long time span. About 500 samples of voltage, current, and active power were collected during normal operation. The data were fed into the proposed system and resulted in 97% accuracy rate in trip fault prediction. Ha and Subramanian [14] presented a novel solution for distribution feeder relays for predicting the faults levels. This technique implemented with two main inputs voltage and current of the breakers. The fault current was calculated using Thevenin's theorem and actual measurement was compared. The output of the NN algorithm showed an accuracy of about 98% with less than 2% error rate. Moreover, Di *et al.* [15] proposed a systematic approach that investigates the fault of power electronics elements under different working conditions. Investors and rectifiers are crucial elements in power conversion. However, the life cycle of these components is influenced by a concurrent number of operations. The authors implemented multiple machine learning techniques that took into account the operation condition and data imbalance for efficient converters fail prediction. Multiple probabilistic models were used such as support vector machines (SVM) and self-organizing map (SOM). The final results showed variance with the best prediction for the ensemble classification.

Mahdi and Genc [16] proposed an artificial system for predicting the power network stability after the fault is cleared. The input variables used are fault statues such as prefault, during-fault, and postfault values. The proposed NN uses the cross-entropy function as the cost function to optimize the weights, and the softmax is used as the activation function. Data were divided into three sets: 60% for training, 20% for validation, and finally, 20% for testing. The results of the simulation have shown an overall accuracy of 99.3%. Xiao and Ai [17] presented a statistical NN approach for predicting the power quality disturbances that may affect the power grid. The authors have used the multihidden Markov model. The dataset of power quality disturbance and weather condition were used as the main data to train the model. Moreover, the authors have used Hadoop clustering to process the data efficiently and to reduce computation time. The authors provided that an improvement of 20% was achieved compared to other model used. Kim *et al.* [18] proposed a predictive NN model to predict and evaluate the dissolved gas analysis in substation transformers based on the previous history of operation. Optimization technique was used to solve the fitting issue. The data were collected from seven substation transformers. Transformer's health status was recorded using the supervisory control and data acquisition

(SCADA) monitoring system. Standard mean absolute error and percentage were used for the regression performance check. After extensive testing, the prediction error of each dissolved gas generated by the increased oil temperature in the transformer index was very low, which was 15% for H_2, 7% for C_2H_2, 5% for C_2H_4, 5% for C_2H_6, and 1.5% for CH_4. The prediction error was limited within 2% for each gas level prediction. The overall prediction accuracy was between 84% and 97%. Lee and Ke [19] presented a monitoring system for large-scale Internet of Things (IoT) in countryside areas. The authors deployed 19 long range (LoRa) nodes over an area with dimensions of 800 m $\times$ 600 m, with access gateway of 1-min interval data collection. The authors provided that the packet delivery ratio (PDR) ratio for the proposed mesh network achieved about 88.49%, whereas the traditional star topology achieved 58.7%. The authors have added that the project aim is to explore the potential of IoT mesh deployment architecture in areas that require long-range transmission. Lv *et al.* [20] proposed optimization clustering method for mixed data for SDN-based smart grid networks. The output algorithm is based on a combination of *k*-means and *k*-modes algorithms. The authors provided that the proposed algorithm satisfies the differential privacy experiment with efficient accuracy. Wang *et al.* [21] adopted an energy efficient sense layers architecture to address the energy that is consumed by large number of IoT nodes. The authors provided that the proposed framework is in three layers, which are sense, gateway, and control layers. The authors used sleep and wake scheduling protocol with prediction of sleep intervals. Furthermore, the authors have deployed in simulation 300 nodes in a large area, where 250 nodes are for sensing and 50 for gateways. After extensive testing, the results show a significant drop in power consumption improving resource utilization and energy consumption. Wang *et al.* [22] discussed concerns of dense deployment of small cells that are inconsistent interfaces, frequent handovers, and extensive backhauling. The authors have introduced SDN for the next generation wireless networks (NWNs) architecture by decomposing the control plane from the data plane. The authors have used virtual radio access technologies (RATs) design to support different services. The authors concluded that the proposed software-defined nework controller (SDNC) is able to predict user's movement path that is near the access point (AP) to implement the handover. After extensive testing, the authors added that the proposed approach was validated and handover is thus accelerated and overall latency is reduced. Wang *et al.* [23] discussed the large amount of data that are generated from big data platforms such as health monitoring networks that require real-time processing and analysis. Many of these data are not needed and cause delay in processing and storage. Therefore, the authors proposed an reduced variable neighborhood search (RVNS) optimization search method that operates in three layers. For that reason, the authors have used three layers approach that are fault-tolerant approach to ensure the reliability of the eHealth system and second is the layer that checks for accuracy of the data and the final layer is where RVNS optimization is implemented where only valuable data will be reported to the health provider system for processing. This approach helps to efficiently increase processing time and delivery ratio. Gao *et al.* [24] proposed multiple approaches starting with a probabilistic modeling using Markov chain method to verify the energy routing system in smart green city networks and markov decision process (MDP) model to check the cost of the service requester and provider. The authors also introduced a monitoring tool over the energy router (ER) system to monitor the scheduling process. The processing of power transactions process was implemented in the cloudlet platform.

The main contributions of this paper can be summarized as follows.

1) We propose a customized SDN infrastructure that consists of long-range power IoT sensor network called grid management network (GMN). The GMN consists of two parts: a) the WSN section that is implemented on each DT per each zone. A list of sensors is installed on each transformer such as a temperature sensor, oil level sensor, humming noise sensor, and overloading sensor. These sensors represent the health status check for each DT. Each one of these sensors is linked to an RF transmitter using long-range LoRa WAN communication. The second part of the architecture is b) wireless-static data center. The data center consists of multiple paths to provide redundancy with fault tolerance route recovery. The network will use SDN controller as the gateway entry node to the static data center.
2) We propose a fault prediction algorithm called distribution transformer fault prediction (DTFP) for fault prediction in DT. The NN algorithm consists of multiple interconnected mesh hidden layers with various weights. Optimization is implemented using back probation to tune the weights for efficient prediction accuracy. The DTFP is installed in the management layer with periodic fault prediction based on hourly historical data.
3) We propose a communication protocol called software-defined communications (SDN-sense) for the wireless IoT nodes on the DTs network. The protocol runs on both layers, control layer represented as sink node and forwarding layer representing the forwarding engines. The forwarding tables are built using received broadcast (BC) packets and then information is relayed to the sink SDN node for constructing the topology table. The best route is selected to be used as the main route; however, an alternative fail recovery is addressed with the most reliable route.

The remainder of this paper is organized as follows. In Section III, system testbed architecture is presented. In Section IV, experimental results and analysis are discussed and explained. In Section V, hardware implementation and deployment are discussed. Finally, Section VI is the conclusion where a summarization of this paper is illustrated.

III. System Testbed Architecture

Our proposed GMN consists of real-virtualized hardware components that run on a Linux server. The core network data center runs on pure SDN architecture that consists of two SDN controllers with a fail-over capability and forwarding engines as OpenFlow vSwitch (2.9.2) [25]. The operating system platform we have used is the Ubuntu Server (14.04.5) [26] with four-port Intel Ethernet NIC cards. We implemented Mininet

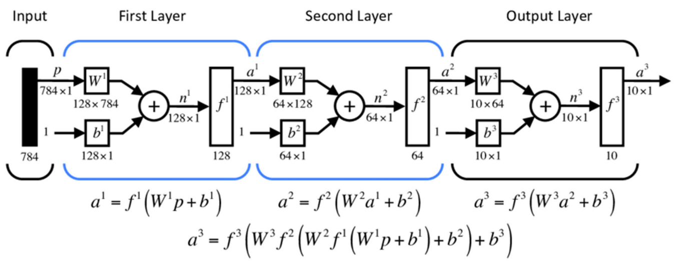


Fig. 1. Multilayer perceptron architecture diagram [29].

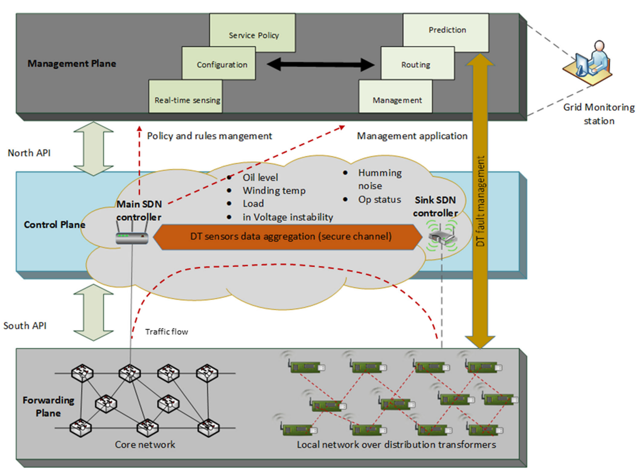


Fig. 2. Proposed architecture block diagram.

[27] and Floodlight Controller [28] as SDN devices that we subsequently modified and developed to match our proposed network architecture requirements and to support the grid network communication. The traffic forwarding depends on the matching statement in the forwarding table. If a match is identified, then the action is to forward the packet to the next device, whereas if no match has been identified, then a query message is sent to the SDN controller (floodlight) to request on what to do to the packet. The SDN controller then responds back by either installing new rules or advising to drop the packet. The wireless sensor nodes are considered one of the crucial parts for the success of this project that provides an essential and precise status overview of the grid. We propose the use of LoRa RF communication that equipped with a variety of sensors to support multifeature sensor readings. We designed the wireless network on the basis of SDN concept, whereas the sink node represents the gateway SDN controller and the rest of the nodes represents OpenFlow forwarding engines. A proposed algorithm that governs the node's communication is defined in Fig. 4.

The experimental setup combines both virtualized and hardware environment, where the virtualized environment represents the core network with the SDN and OpenFlow switches and the hardware setup consists of IoT module with different sensors that are attached on the DT. The sensors measure different parameters that are considered performance degradation factors in the life cycle of a transformer that are as follows: first is the temperature sensor that is installed on the outer tank shell of the transformer, whose output is an analog that is fed to the microcontroller such as Arduino Uno for analog-to-digital conversion

Algorithm 1 FP-sense.

1: **Given** $(x_1, x_2,..., x_n)$, $(y_1, y_2, ..., y_m)$, where $x_i \in X$, $y_m \in Y = \{1, 2, 3, ..., N\}$ ▷ input and output data for grid parameters
2: **Initialize** random values for W_k *and* b_k: $W_k = \sum_{k=1}^{x_n} random.randn(\theta)$ $b_k = \sum_{k=1}^{x_n} random.randn(\beta)$
3: **Define** activation function: $\text{ReLu(z)} = \begin{cases} z, & for\ \ z \geq 0 \\ 0, & for\ \ z < 0 \end{cases}$
4: **Define** model: $\text{Mod}(x_n, W_k, b_k) = \sum[\text{x}_n\ \text{W}_k] + b_k$
5: **For** t in range (1000): **do**
6: **Initiate** $\lambda = 0.01$ ▷ learning rate
7: random = random.randint(arrary[x_n, y_m]) ▷ select random value from array
8: test = array [random]
9: Mod $= \sum test_{1,...,l}.\text{W}_k + b_k$ ▷ calculating the mod value for each entry
10: pred = $\text{ReLu}(\underbrace{Mod(\text{x,w,b})}_{z})$ ▷ initial prediction values
11: $ObjFunc = (\text{pred - } T_k)^2$▷ error value for each target T_k
12: **If** ObjFunc $\geq \psi$ **do** *tune weights* : ▷ checking for error capacity
13: $\dfrac{\text{d}ObjFunc}{\text{d}pred} = 2 \times (pred{-}\text{T}_k)$ ▷ starting backpropagation to use with tuning W_k and b_k
14: $\dfrac{\text{d}pred}{\text{d}z} = drvReLu(z) \times (pred{-}\text{T}_k)$
15: $\dfrac{\text{d}z}{\text{d}W_k} = test[l]$
16: $\dfrac{\text{d}z}{\text{d}b_k} = 1$
17: $\dfrac{\text{d}ObjFunc}{\text{d}W_k} = \dfrac{\text{d}ObjFunc}{\text{d}pred} \times \dfrac{\text{d}pred}{\text{d}z} \times \dfrac{\text{d}z}{\text{d}W_k}$ ▷ calculating change in error with regards to W_k
18: $\dfrac{\text{d}ObjFunc}{\text{d}b_k} = \dfrac{\text{d}ObjFunc}{\text{d}pred} \times \dfrac{\text{d}pred}{\text{d}z} \times \dfrac{\text{d}z}{\text{d}b_k}$ ▷ calculating change with respect to b_k
19: **Calculate** new weights and biases:
20: $W_{k_{new}} = W_k \times \lambda \times dObjFunc/\text{d}W_k$ ▷ new values of W and b to decrease the error in pred
21: $b_{k_{new}} = b_k \times \lambda \times dObjFunc/\text{d}b_k$
22: **End If**
23: **End For**
24: **Check** for error level:
If Objfunc ¡ ψ **then** :
FaultPred = pred(z) ▷ predicted value of the local DT fault
3
25: **Else: Go** to **step 4** ▷ repeat iteration with fixed weights
26: **End If**

Fig. 3. Proposed DTFP algorithm pseudocode.

and then to the LoRa module for transmission; and second is the oil level sensor that is placed inside the oil tank to measure the decreased oil levels. The output of the analog voltage is supplied to the microcontroller for conversion to readable value. The overloading profile monitoring is read using a sensor that measures voltage, current, and power factor. The last sensor used is the humming noise. Many transformers in Iraq suffer from

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
CONDITION SI

Status Index (%)	Condition
100<SI<90	very good
80<SI<70	good
70<SI<65	yellow alert (require investigation)
SI<60	system critical (fail)

TABLE II
INVESTIGATED TRANSFORMER SPECIFICATION

Parameter	Description
Rated voltage(max)	11kVA
Rated voltage (low)	433v-250v
Load current (max)	3.3A
Load current (high)	84A
Connection	Delta
No. of phases	3
Frequency	50 c/s
Noise level	50db
Operating average temperature	35-40 Deg.C

the noise instability that is important to be measured to provide preventive maintenance if required.

Respectively, the next main part of our architecture is to provide fault prediction over transformer operational cycle. The model we proposed is able to predict the faults based on the previous historical data of the sensors. Fig. 1 shows a typical NN model with multiple hidden layers.

The main factors in calculating the error level in our prediction model and to test the usefulness of our fault prediction platform are the mean square error (MSE) and root-mean-square error (RMSE), both of which typically are called objective or cost function. The cost has to be a very small value in order for our system to be reliable in fault prediction analysis. The MSE and RMSE can be expressed as in (1) and (2). The difference between the two equations is that taking the RMSE gives high weights to large errors, which can be exceptionally useful when undesirable errors occur

$$\mathrm{Obj}(x_1, x_2, \ldots, x_n) = \frac{1}{n}\sum_{i=1}^{n}(\mathrm{Flt}_{\mathrm{pred}} - \mathrm{Flt}_{\mathrm{trgt}})^2 \tag{1}$$

$$\mathrm{Obj}(x_1, x_2, \ldots, x_n) = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}(\mathrm{Flt}_{\mathrm{pred}} - \mathrm{Flt}_{\mathrm{trgt}})^2}. \tag{2}$$

The block diagram of the proposed architecture is depicted in Fig. 2. The diagram consists of a forwarding plane that represents the OpenFlow engines for the core and wireless network. The next layer is the control plane that represents the data steering and route management platforms. A secure channel is established between the main SDN and sink SDN controller for data security and reliability. Furthermore, the top layer is the management layer that represents the storage and real-time sensing platform with the fault prediction network that is trained repeatedly with new events every hour. The retraining is implemented using the backpropagation model. After using our developed proposed DTFP algorithm as in Fig. 3, we found the proposed model achieves efficient accuracy in prediction of faults.

Algorithm 1 SDN-sense.

```
Initialize DisSink=0; rpkt=false; Nirt=0;
    Tmr=k, where k ∈ range (20 − 30)
Initiate SDN: {                          ▷ at the SDN level
Braodcast DiscPkt:
Nirt=Tmr                  ▷ Max time to wait for response
Listen to QueryReply from NeiNode
}
Receive DiscPkt              ▷ for each n node ∈ cluster
If (rpkt == false) {
ftable = DiscPkt            ▷ building forwarding table
rpkt = true
DisSink+=1
Broadcast DiscPkt
Listen to QueryReply_local
Nirt=Tmr
}
Else If Nirt == 0 {          ▷ no response; timer reset
Braodcast FeedBpkt             ▷ repeat for each node
    ∈ cluster
rpkt=false
DisSink-=1
Update ftable          ▷ updating forwarding table with
    lower node info
Broadcast join-query → SDN_main
Listen join-conf pkt, Tmr=k ▷ conf from SDN main
    to start sending pkts
If Tmr==0; Go to 21
Forward data[i] → SDN_main                 ▷ start data
    aggregation to core network
}
Else If DisSink==0 {        ▷ reached SDN controller
Install FeedBpkt in Ttable ▷ topology table update
DisNode = Filter(DisSink) ▷ select shortest path to
    each n Node
}
If DiscPkt ==NULL {                        ▷ failover case
Nirt=Tmr
node1 = NewSDN              ▷ announce new SDN with
    DisSink=1
Go to step 2
}
End
```

Fig. 4. Proposed SDN-sense pseudocode.

In order to track the overall status of the DT system, we assume a status index (SI) factor of the DT that is considered a powerful tool for identifying the overall operational health status of the system. We assume that the SI is based on scale (0–1) where 0 is no critical status and 1 is a high critical health condition, whereas the subdivisions between 0 and 1 are considered the real operation values of the transformer. If we log the sensor data into a sigmoid function, we can get the probability of how well the transformer is performing. Let x_i be a variable SI that represents the status of the specific sensor. The SI for multivariable inputs can be expressed in a logistic regression model as follows:

$$\mathrm{SI}(\%) = \frac{1}{1 + e^{-\sum_{i=1}^{n}(\alpha x_i)}} \times 100 \tag{3}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

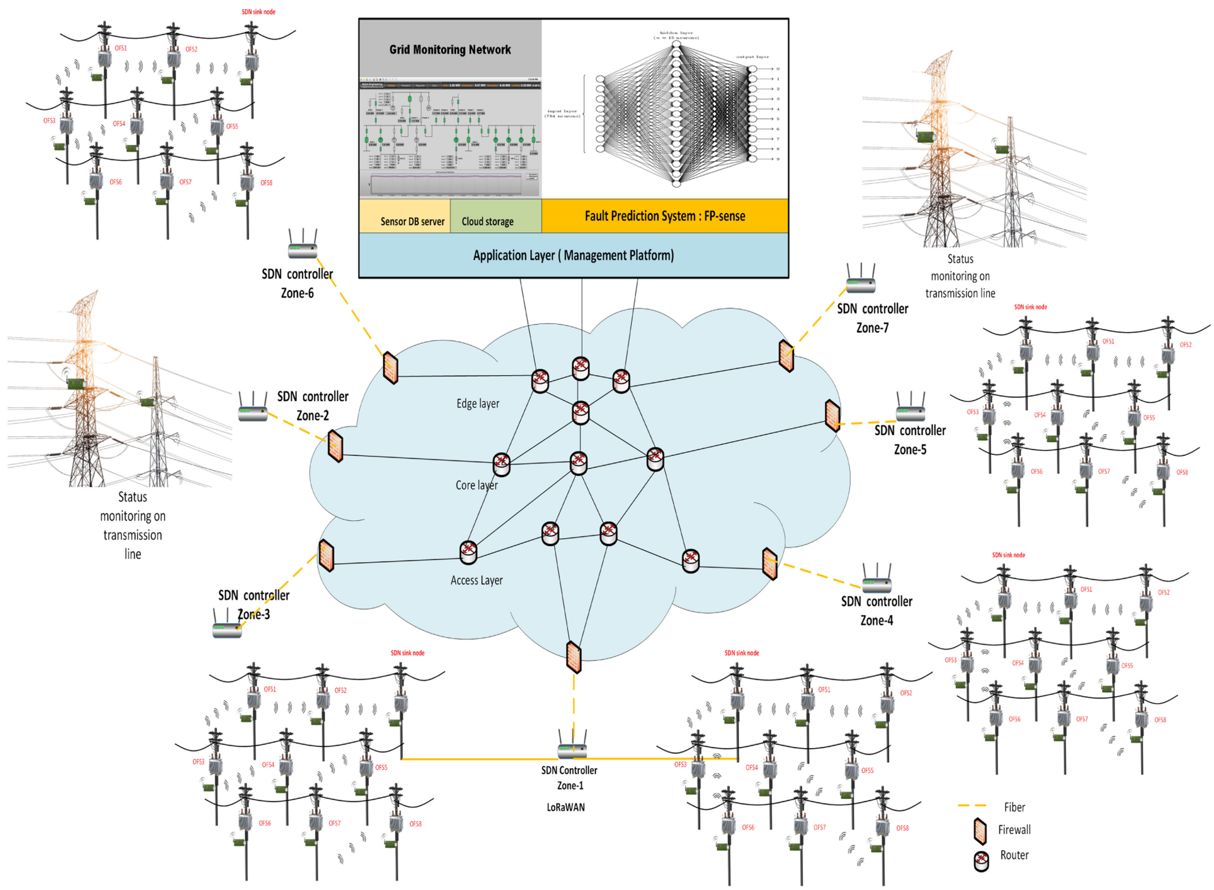


Fig. 5. GMN proposed architecture platform.

where α represents the weight effect of each sensor variable that ranges between 1 and 10. We can classify the SI index of the DT health status as shown in Table I.

The proposed communication algorithm between the IoT nodes is governed using SDN-sense algorithm, as described in Fig. 4. The DT that is being investigated is described in Table II.

IV. Experimental Results and Analysis

The proposed complete smart grid architecture-based SDN is described in Fig. 5. We can notice that the core network that is represented as the cloud consists of multipath routing links governed by SDN enabling forwarding engines to operate their designated operating requests. Furthermore, each specified zone is connected via a mesh network of SDN and OpenFlow switches that relay sensor data to the main SDN controller, which is installed at the edge of the cloud. The cloud that we used in our experimental testing is based on virtualized environment represented in virtual machines. The SDN controller is implemented using Mininet and Floodlight Controller. The rules were configured in advance and set to be installed in the Open vSwitch engines.

The wireless nodes communication is based on our proposed algorithm SDN-sense, as shown in Fig. 4. Moreover, the architecture described in Fig. 5 represents the overall proposed architecture that combines the core network represented in the cloud, the SDN wireless mesh network, and the fault prediction system. We consider the SDN architecture as a directed graph $G = (\mathrm{SW}, L)$, where SW represents all the OpenFlow switches (SDN to be in case of failure), and L represents a set of RF links $L = \{(i,j) \in S \times S, i \neq j\}$. The SDN and open flow (OF) switches are customized and programmed to match our case study, and the SDN can be accessed via a Python application programmable interface (API) for further modification and data retrieval. Furthermore, the OF forwarding table can be represented as $F = \{\lambda_{\text{pkt}}, \beta_{\text{tab}}, \alpha_{\text{act}}\}$, where the three main objects compose the forwarding table, which are flows, tables, and actions. Each packet is required to be matched to a table and then a decision is made on where to forward the packet based on a bucket of actions. The number of rules that is existed in a particular OF node can be represented as follows:

$$R_k = \sum_{1}^{n} \Delta_{k,t} \tag{4}$$

where $\Delta_{k,t}$ represents the rule per OF node with t as an indication for subrule. k is subscript of total rules. The total matching delay that may occur in the OpenFlow table can be denoted as follows:

$$\phi_{\text{match-delay}} = \sum_{1}^{N} R_k \times \sigma_{q\text{-factor}} \tag{5}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

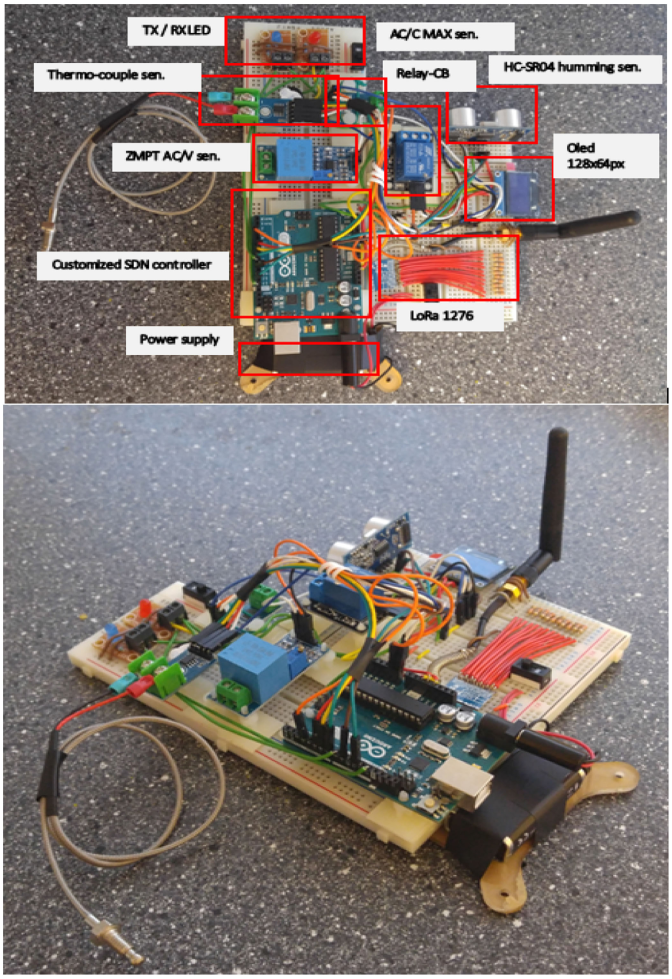


Fig. 6. Proposed OpenFlow IoT sensor platform.

where $\sigma_{q\text{-factor}}$ is the queueing delay for processing flows that can affect the total processing capacity of the OF node significantly. Although power consumption of the SDN sink node is not high, it is worthy to mention it as it may affect on the lifetime of the node sensor and designing an efficient power management node can result in efficient power consumption and longevity of the operating node. The power consumption of the SDN sink and main SDN node can be expressed in (6) and (7)

$$P_{\text{sink}_{\text{total}}} = \sum_{1}^{n} \theta_{\text{temp}} + \sum_{1}^{n} \theta_{\text{oil}} + \sum_{1}^{n} \theta_{\text{temp}} + \sum_{1}^{n} \theta_{C-\text{in}} + \sum_{1}^{n} \theta_{V-\text{in}} + \sum_{1}^{n} \theta_{\text{lora}} \quad (6)$$

where θ represents the inbound traffic power consumption of a specific sensor

$$P_{\text{SDN}_{\text{main}}} = \sum \lambda_{\text{clust}_1} + \sum \lambda_{\text{clust}_2} + \sum \lambda_{\text{clust}_3} + \cdots + \sum \lambda_{\text{clust}_n} + \sum \lambda_{\text{lora}} + \sum \lambda_{\text{init}}. \quad (7)$$

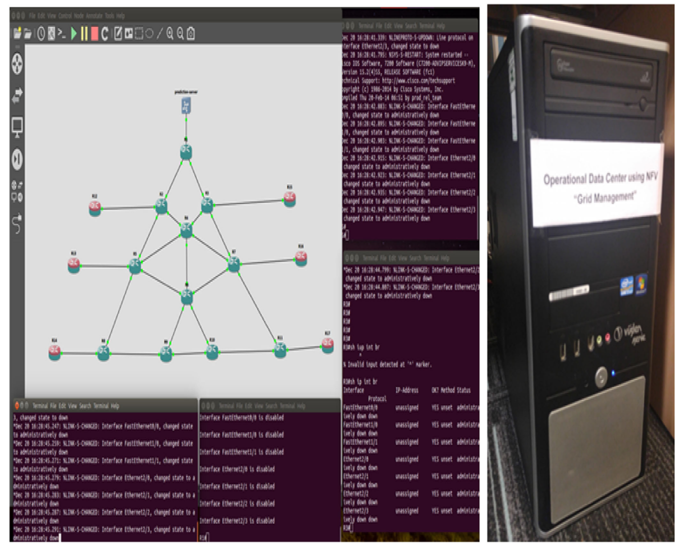


Fig. 7. Proposed virtualized core topology implemented on a Linux server for sensor traffic management and fault prediction using Python classification libraries.

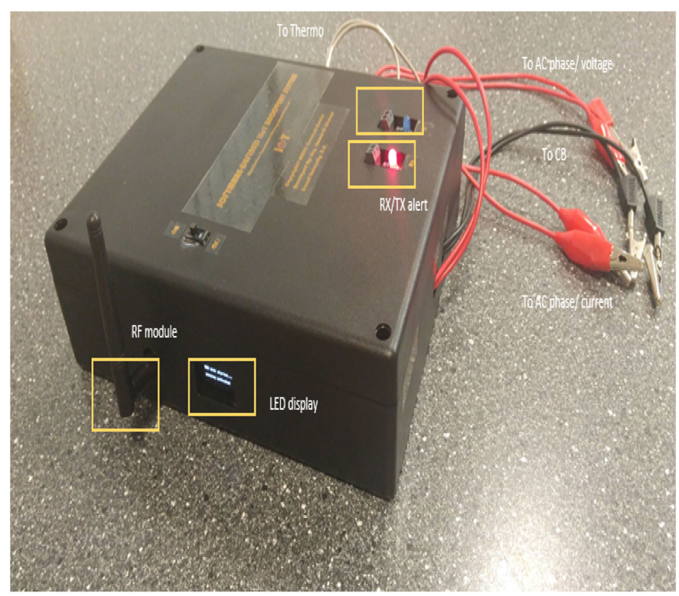


Fig. 8. Finalized proposed sensor platform while in active mode.

V. Hardware Implementation and Deployment

In this section, we present the proposed sensor hardware that can be implemented in a residential transformer zone. The system is built using an IoT off-the-shelf hardware that is programmed with SDN implementation principle. The hardware unit of the sensor consists of an Arduino board that is programmed as a microcontroller board with SDN functionality. The proposed hardware consists of five main sensors that are temperature sensor, oil level sensor, humming noise sensor, ac-in sensor, and V-in sensor. The sensor nodes based on OpenFlow platform communicate with sink SDN node using a long-range communication network by implementing LoRa network due to the heterogeneity of the communication in such environment.

Fig. 9. Faulty DT samples from Baghdad electrical grid maintenance site. From left to right: pivot damage, oil leaks, and shorted winding.

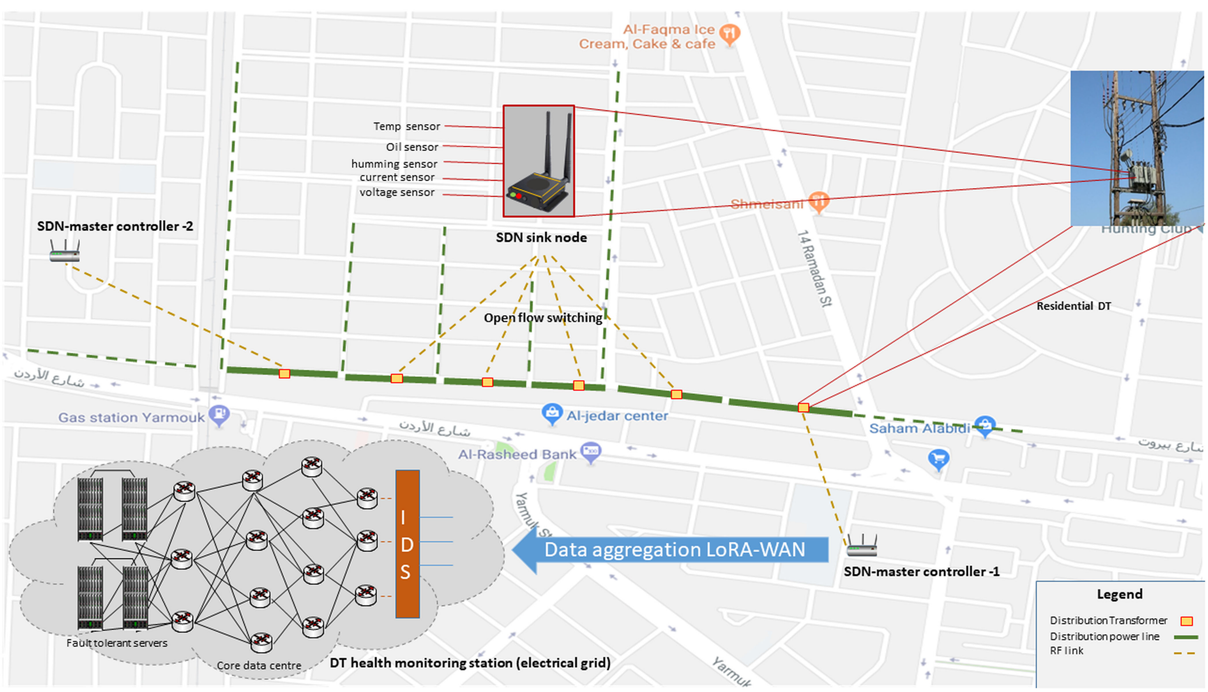


Fig. 10. Proposed SDN sensor deployment scenario in a residential zone per DT platform.

The main gateway or SDN main responsible for managing the communication with all sink nodes and to collect all sensor data to be forwarded for aggregation to the data center. After data are processed and stored, they will be fed to the prediction system so that a fault prediction can be produced based on real-time sensor data. The prediction can help in identifying any future faults that could occur in the D-Transformer and to reroute power and isolate faulty D-transformer for a maintenance procedure. Fig. 6 shows the proposed SDN IoT hardware prototype with components labeled, whereas Fig. 7 shows the virtualized data center implementation that runs on a Linux server. The server represents the core network that uses network function virtualization for efficient power consumption management. The prediction system runs on the servers under Python library. In our IoT testbed, we have used Arduino Uno [30] and programmed it as a customized SDN sink controller that operates with a LoRa module LX1278 with a custom-tailored antenna for better signal gain and propagation. The rest of the sensors are ac sensor ACS712 with ac voltage sensor ZMP101B. For the temperature sensor, we have used thermocouple sensor MAX6675 with an ultrasonic sensor to measure the humming noise HCSR04. The sink node is powered with 9 V power supply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

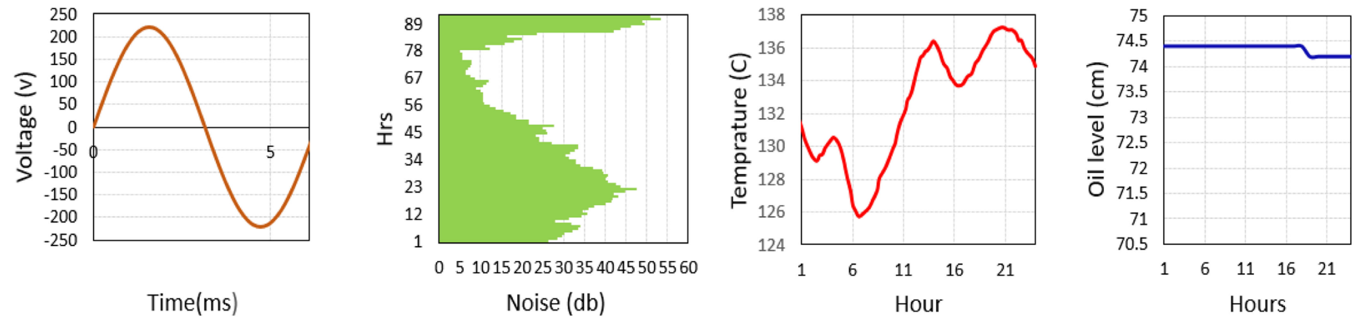


Fig. 11. Sensor data collected for one phase from an operational miniature DT using our proposed testbed.

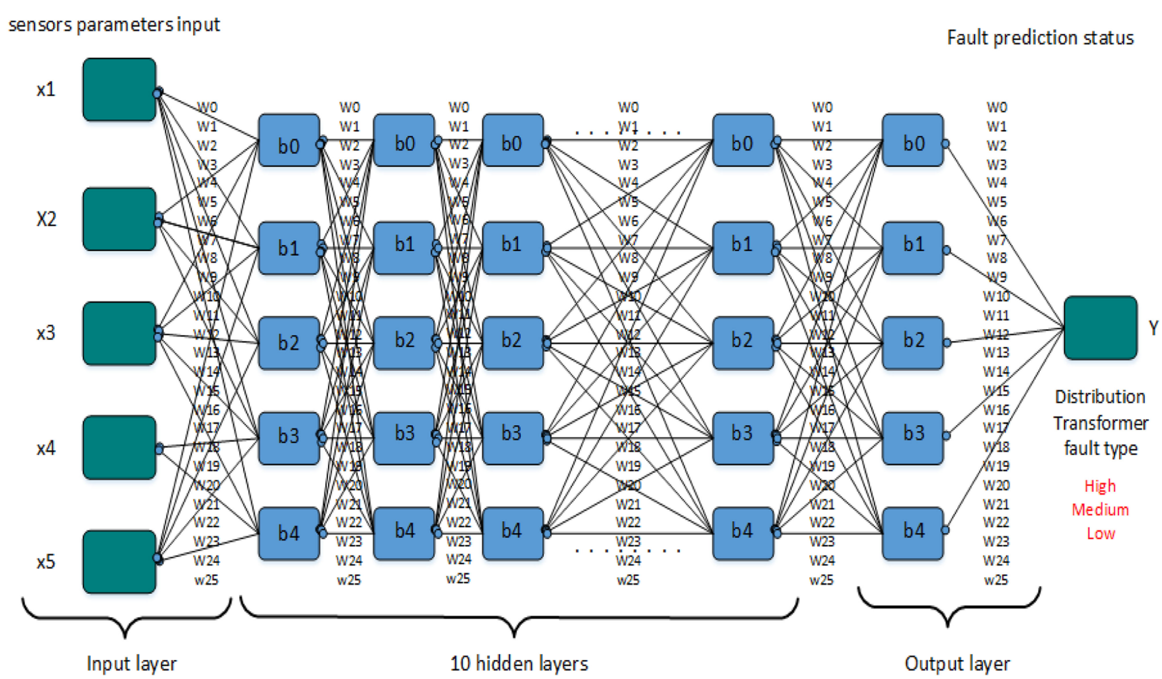


Fig. 12. Proposed combined feedforward and decision tree fault classification system structure.

Fig. 8 shows the final enclosure box for the proposed testbed. This box is designed for a single phase only for this current research project. The three main cables are for ac-V, ac-C, and circuit breaker (CB) for switching and transformer protection.

Fig. 9 shows faulty transformers images that were collected from different grid sites in Iraq that were effected by many factors such as shorted winding, high-temperature fault, high incoming voltage, and oil leaks. Additionally, damages could be caused due pivot pole fall, which causes total damage to the D-transformer outer case.

In Fig. 10, we present a deployment case scenario of our proposed SDN sink sensor over residential transformers. The sink node communicates to the SDN master node using LoRa RF communication and then to the cloud network for sensor data processing. The IoT-based sensor node is based on sensing and action implementation based on the level of incoming data from each sensor. Many sensors have been implemented in our testbed such as oil level sensor, temperature sensor, ac voltage sensor, and ac current sensor. Based on these data, an action will be made to cut off the circuit breaker in case of a high alert. Additionally, these data will be fed to the prediction system for statistical analysis based on real-time data and historical environmental data. A decision will be generated from the prediction system to regulate the transformer behavior and to reduce any future fails. The testbed prototype that we have implemented only suits single case transformer scenario. However, it operates as a testbed that can operate with three-phase system. The testing was implemented on a small miniature scale transformer due to limited resources. Fig. 11 shows sensor readings for a single-phase transformer that depicts the health index of each part.

In order to build our proposed prediction system, we have used historical dataset for the outages and faults based on records that were logged by the grid transformer maintenance workshop in Iraq. The dataset includes data such as fault time, fault date, fault type, and fault number of occurrences. These attributes

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

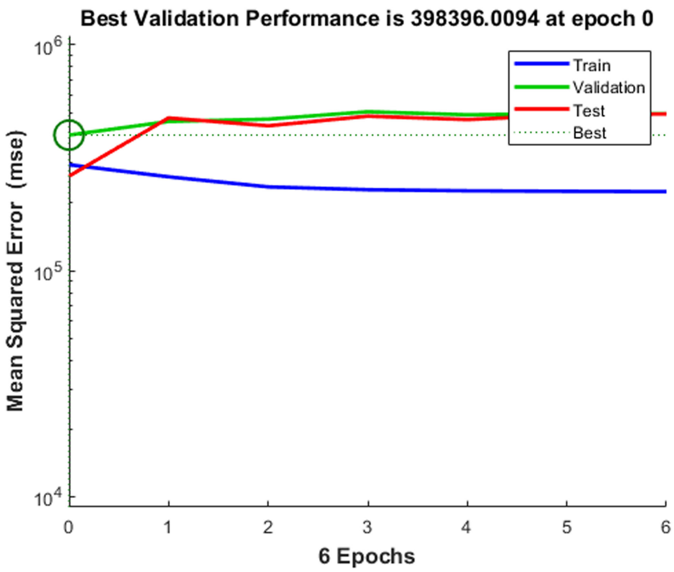


Fig. 13. Objective function error rate.

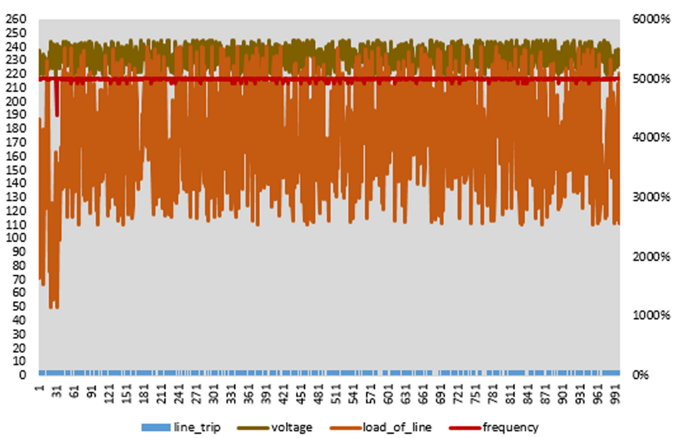


Fig. 14. Dataset used for training the network.

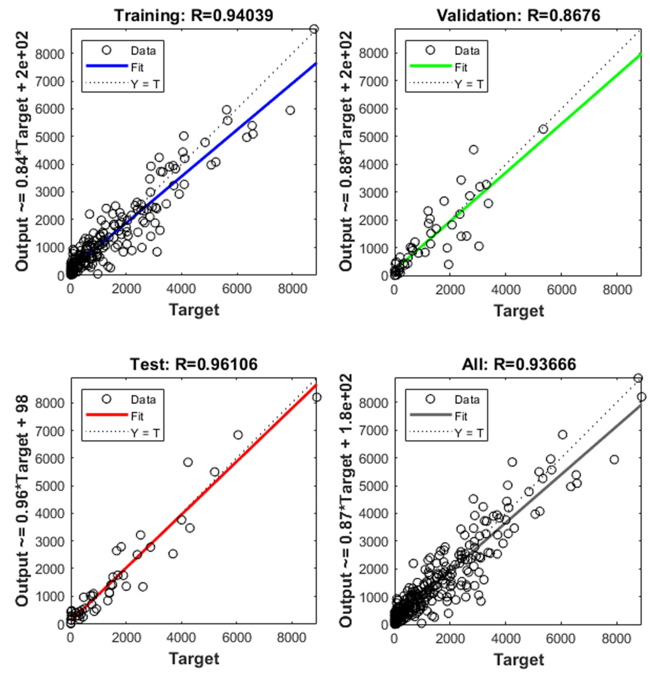


Fig. 15. Prediction phases.

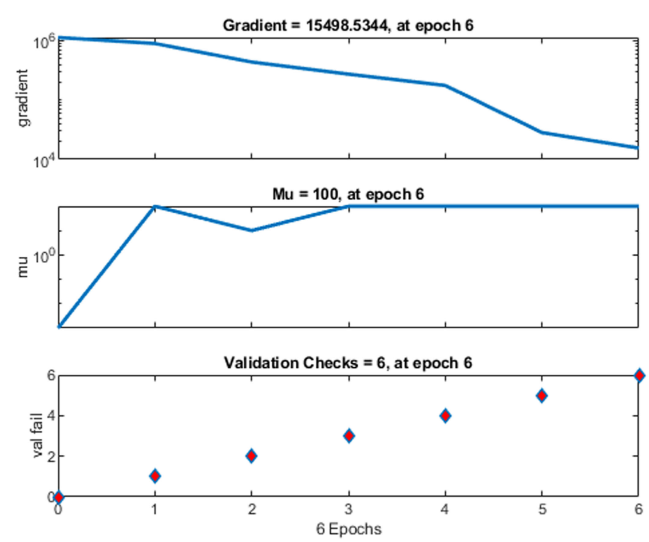


Fig. 16. Gradient and validation checks.

were taken as input to the NN with an additional sensor data that can increase the accuracy in a form of double stage input. The hidden layer as we see in Fig. 12 consists of ten layers with the sigmoid as the objective function.

The weights were initialized randomly at first stage and then backpropagation was used to tune the weights for better prediction accuracy. However, for final stage prediction, we have used decision tree classification algorithm for accurate prediction. We have implemented a combination of feedforward for sensors and historical dataset and decision tree algorithm for finding the best average prediction of historical and real combined sensor data output. We can notice that a better accuracy has been achieved by using our proposed work of feedforwarding and decision tree averaging algorithm while minimizing the error rate between each Y prediction value. Furthermore, the sensor data were fed to our proposed prediction platform and we were able to get a low error rate after 1000 rounds of training, as depicted in Fig. 13. The optimization of the error rate can be reduced more by using more critical relational parameters that can be estimated for each transformer.

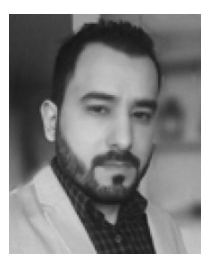
Fig. 14 represents the dataset parameters that were used to train our proposed model. The main input data are the line trip, frequency, line load, and voltage. Fig. 15 shows the prediction of the type of fault and phase line overload with 96.1% accuracy. The accuracy can be optimized more by using more operational parameters. Moreover, Fig. 16 shows the gradient decent process parameters that are used to tune the weights to minimize the cost function.

VI. Conclusion

The current electrical grid system in Iraq needs to be updated with new engineering implementation to overcome demand and outage challenges and adapt itself to new grid requirements to reduce maintenance cost. Therefore, this paper proposed a novel SDN IoT sensor platform to monitor the electrical parameters in DTs to provide solutions for the electrical grid in Iraq. Current electrical grid weaknesses were discussed and the effectiveness of our proposed system was highlighted along with the proposed prediction system. Experimental testing was implemented on an application case to validate the proposed prototype. The hardware was built using IoT hardware sensors and controllers. The controller was programmed as a customized SDN controller with the ability to operate as a sink and regular node. The testbed can also be connected to a circuit breaker to smartly manage any high alert threshold that could occur, such as overload, high voltage, etc. The SDN-sense protocol was proposed to manage the communication of N nodes efficiently. Moreover, we have implemented the data center on a virtual Linux server with multiple paths for redundancy. The prediction platform was implemented using a Python library and MATLAB simulation. Experimental results showed prediction results with 96.1% accuracy. In summary, the proposed system is considered to be a low-cost implementation with real-time management that can provide a total overview of the DT status and eliminate any future fails and outages that may occur in the distribution lines.

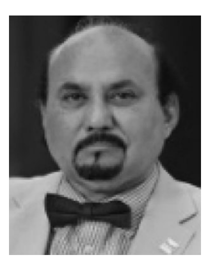
REFERENCES

[1] M. Wang, A. J. Vandermaar, and K. D. Srivastava,“ Review of condition assessment of power transformers in service,” *IEEE Elect. Insul. Mag.*, vol. 18, no. 6, pp. 12–25, Nov./Dec. 2002.
[2] K. Wang *et al.*, “A survey on energy internet: Architecture, approach, and emerging technologies,” *IEEE Syst. J.*, vol. 12, no. 3, pp. 2403–2416, Sep. 2018.
[3] S. Azodolmolky, *Software Defined Networking With OpenFlow*, vol. 21. Birmingham, U.K.: Packt Publ., 2003.
[4] E. C. Fontana, “Powering software defined networks with an any rack capable power bus,” in *Proc. IEEE Int. Telecommun. Energy Conf.*, 2016, pp. 1–4.
[5] T. D. Nadeau and K. Gray, *SDN: Software Defined Networks: An Authoritative Review of Network Programmability Technologies*. Newton, MA, USA: O’Reilly Media, Inc., 2013.
[6] S. Bailey *et al.*, “SDN architecture overview,” Open Networking Foundation, Menlo Park, CA, USA, 2013. [Online]. Available: https://www.opennetworking.org/images/stories/downloads/sdn-resources/technical-reports/SDN-architecture-overview-1.0.pdf
[7] J. Li, D. Li, Y. Yu, Y. Huang, J. Zhu, and J. Geng, “Towards full virtualization of SDN infrastructure,” *Comput. Netw.*, vol. 143, pp. 1–14, 2018. [Online]. Available: https://doi.org/10.1016/j.comnet.2018.06.014
[8] N. Bizanis and F. A. Kuipers, “SDN and virtualization solutions for the internet of things: A survey,” *IEEE Access*, vol. 4, pp. 5591–5606, 2016.
[9] International Electrotechnical Commission, “Internet of Things: Wireless sensor networks,” IEC White Paper, 2018. [Online]. Available: http://www.iec.ch
[10] I. N. da Silva *et al.*, “Intelligent systems for the detection of internal faults in power transmission transformers,” in *Advances in Expert Systems*. Rijeka, Croatia: InTech, 2012.
[11] E. Alpaydin, *Introduction to Machine Learning*. Cambridge, MA, USA: MIT Press, 2014.
[12] R. Amin, M. Reisslein, and N. Shah, “Hybrid SDN networks: A survey of existing approaches,” *IEEE Commun. Surv. Tut.*, vol. 20, no. 4, pp. 3259–3306, Oct.–Dec. 2018.
[13] S. Zhang, Y. Wang, M. Liu, and Z. Bao, “Data-based line trip fault prediction in power systems using LSTM networks and SVM,” *IEEE Access*, vol. 6, pp. 7675–7686, 2018.
[14] H. Ha and S. Subramanian, “Predicting the prospective fault level on distribution grids and its impact on protective relaying,” in *Proc. 70th Annu. Conf. Protective Relay Eng.*, 2017, pp. 1–4.
[15] Y. Di *et al.*, “Fault prediction of power electronics modules and systems under complex working conditions,” *Comput. Ind.*, vol. 97, pp. 1–9, 2018.
[16] M. Mahdi and V. I. Genc, “Artificial neural network based algorithm for early prediction of transient stability using wide area measurements,” in *Proc. 5th Int. Istanbul Smart Grid Cities Congr. Fair*, 2017, pp. 17–21.
[17] F. Xiao and Q. Ai, “Data-driven multi-hidden Markov model-based power quality disturbance prediction that incorporates weather conditions,” *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 402–412, Jan. 2019.
[18] Y. Kim, A. Goyal, and T. Kumar, “Predictive modeling of dissolved gas concentration in oil-immersed substation transformers,” in *Proc. IEEE Smart Energy Grid Eng.*, 2016, pp. 261–267.
[19] H.-C. Lee and K.-H. Ke, “Monitoring of large-area IoT sensors using a LoRa wireless mesh network system: Design and evaluation,” *IEEE Trans. Instrum. Meas.*, vol. 67, no. 9, pp. 2177–2187, Sep. 2018.
[20] Z. Lv *et al.*, “An optimizing and differentially private clustering algorithm for mixed data in SDN-based smart grid,” *IEEE Access*, vol. 7, pp. 45773–45782, 2019.
[21] K. Wang, Y. Wang, Y. Sun, S. Guo, and J. Wu, “Green industrial internet of things architecture: An energy-efficient perspective,” *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 48–54, Dec. 2016.
[22] K. Wang, Y. Wang, D. Zeng, and S. Guo, “An SDN-based architecture for next-generation wireless networks,” *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 25–31, Feb. 2017.
[23] K. Wang, Y. Shao, L. Shu, C. Zhu, and Y. Zhang, “Mobile big data fault-tolerant processing for ehealth networks,” *IEEE Netw.*, vol. 30, no. 1, pp. 36–42, Jan./Feb. 2016.
[24] M. Gao, K. Wang, and L. He, “Probabilistic model checking and scheduling implementation of an energy router system in energy internet for green cities,” *IEEE Trans. Ind. Inform.*, vol. 14, no. 4, pp. 1501–1510, Apr. 2018.
[25] Linux Foundation Projects, “OPNFV,” 2018. [Online]. Available: https://www.
opnfv.org/about
[26] Ubuntu 14.04.5 LTS (Trusty Tahr), 2018. [Online]. Available: http://releases.ubuntu.com/14.04/. Accessed on: Oct. 8, 2014.
[27] Mininet, 2018. [Online]. Available: http://mininet.org/. Accessed on: Oct. 8, 2014.
[28] Floodlight Controller, 2018. [Online]. Available: http://www.projectfloodlight.org/floodlight/. Accessed on: Oct. 8, 2014.
[29] G. A. Valencia-Zapata, D. Mejia, G. Klimeck, M. Zentner, and O. Ersoy, “A statistical approach to increase classification accuracy in supervised learning algorithms,” *IPSI BgD Trans. Internet Res. J.*, vol. 13, no. 2, Jul. 2017. [Online]. Available: http://ipsitransactions.org/journals/papers/tir/2017july/p4.php
[30] Arduino, “Ardunio board,” 2018. [Online]. Available: https://www.arduino.cc/



Ammar K. Al Mhdawi (M’18) is currently working toward the Ph.D. degree in electronic and computer engineering with Brunel University London, Uxbridge, U.K.

He is also a Cisco Certified Professional Engineer with more than ten years of experience in network and communications engineering. His research interests include IoT, IIoT, software-defined networks, wireless sensor networks, network function virtualization, delayed tolerant networks, wind energy, and smart grid.



Hamed Saffa Al-Raweshidy (SM’05) received the B.Eng. and M.Sc. degrees from the University of Technology, Baghdad, Iraq, in 1977 and 1980, respectively, the Post Graduate Diploma degree from Glasgow University, Glasgow, U.K., in 1987, and the Ph.D. degree from the University of Strathclyde, Glasgow, U.K., in 1991.

He was with the Space and Astronomy Research Center, Baghdad, Iraq; PerkinElmer, Inc., Waltham, MA, USA; British Telecom, London, U.K.; Oxford University, Oxford, U.K.; Manchester Metropolitan University, Manchester, U.K.; and Kent University, Canterbury, U.K. He is currently the Director of the Wireless Network Communications Center, Brunel University London, Uxbridge, U.K.