Tanzania Journal of Science 47(3): 1138-1153, 2021
ISSN 0856-1761, e-ISSN 2507-7961

Maintenance Automation Architecture and Electrical Equipment Fault Prediction Method in Tanzania Secondary Distribution Networks

Hadija Mbembati, Kwame Ibwe* and Baraka Maiseli
Department of Electronics and Telecommunications Engineering, College of Information and Communication Technologies, University of Dar as Salaam,
P. O. Box 33335, Dar as Salaam, Tanzania
**Corresponding author. E-mail addresses: kwame.ibwe@udsm.ac.tz, hadija@udsm.ac.tz, barakaezra@udsm.ac.tz*
Received 28 Apr 2021, Revised 24 Jul 2021, Accepted 28 Jul 2021, Published Aug 2021
DOI: https://dx.doi.org/10.4314/tjs.v47i3.23

Abstract
Distribution networks remain the most maintenance-intensive parts of power systems. The implementation of maintenance automation and prediction of equipment fault can enhance system reliability while reducing the overall costs. In Tanzania, however, maintenance automation has not been deployed in secondary distribution networks (SDNs). Instead, traditional methods are used for condition prediction and fault identification of power assets (transformers and power lines). These (manual) methods are costly and time-consuming, and may introduce human-related errors. Motivated by these challenges, this work introduces maintenance automation into the network architecture by implementing effective maintenance and fault identification methods. The proposed method adopts machine learning techniques to develop a novel system architecture for maintenance automation in the SDN. Experimental results showed that different transformer prediction methods, namely support vector machine, kernel support vector machine, and multi-layer artificial neural network, give performance values of 96.72%, 97.50%, and 97.53%, respectively. Furthermore, oil based performance analysis was done to compare the existing methods with the proposed method. Simulation results showed that the proposed method can accurately identify up to ten transformer abnormalities. These results suggest that the proposed system may be integrated into a maintenance scheduling platform to reduce unplanned maintenance outages and human maintenance-related errors.

Keywords: Predictive maintenance; fault identification; fault prediction; maintenance automation; secondary electrical distribution network.

Introduction
Secondary electrical distribution network (SDN) is considered as a means for transporting electrical energy in power networks. SDN is divided into two parts, namely primary and secondary distribution networks. The factor that distinguishes these two networks is the network capacity: 33 kV/11 kV for the primary distribution network, and 400 V/230 V for the secondary distribution network. The role of SDN is to supply electrical power directly to low voltage users. In recent years, Tanzania has been in the process of reforming the electric power system. This reformation has significant impacts on the investment and construction of the power distribution networks. Therefore, it seems important to focus on the management of the operation, maintenance, and fault prediction of equipment as an attempt to generate reliable

power supply to the end users (Kalinga et al. 2017).

Population growth and technological advancement have led to the rapid expansion of SDN. The advances in technology have transformed a traditional SDN into a smart electrical grid, which deploys a two-way communication to enhance power reliability, safety, and quality of service (Dileep 2020). The two-way communication system in the smart electrical grid supports several tools, including a self-monitoring tool. This advantage allows the utility company to apply more advanced maintenance strategies to ensure reliability of the electric grid while minimizing human interventions (Colak et al. 2016).

SDN expansion creates higher chances of the occurrence of electrical stress faults in the network. This challenge may lead to unexpected failure of the equipment—a consequence that may generate safety issues and unnecessary cost to the utility companies and consumers. The equipment that shows undesirable conditions can be monitored for prediction and maintenance scheduling. In most SDNs, especially those of Tanzania, equipment monitoring and maintenance scheduling are performed manually due to lack of efficient automation systems that can continuously monitor equipment parameters to identify failure modes (Mnyanghwalo et al. 2019). Traditional preventive maintenance, that uses manual statistical modeling approaches, is still performed frequently by engineers and technicians. Consequently, failure mode and causes analysis may take a long time, and may be susceptible to high probability of human-related errors. This manual process imposes losses to end-users as well as to utilities because such errors introduce additional costs to the companies (Asadzadeh and Azadeh 2014). Currently, SDN lacks automation for maintenance decision support in remotely monitored distribution equipment. Online monitoring of transformer through dissolved gas analysis has been done at the grid substation 60 MVA, 220/33 kV Mwakitebe, Mbeya region (TANESCO). Usually, the high voltage network is fully automated with reliable communication infrastructure and remotely controllable devices. Compared with high voltage network, the SDN is experiencing a lot of stress faults that decrease the lifespan of transformers (Bhargava et al. 2020); examples of these stress faults are tripping, overcurrent, and earth fault (Figure 1).

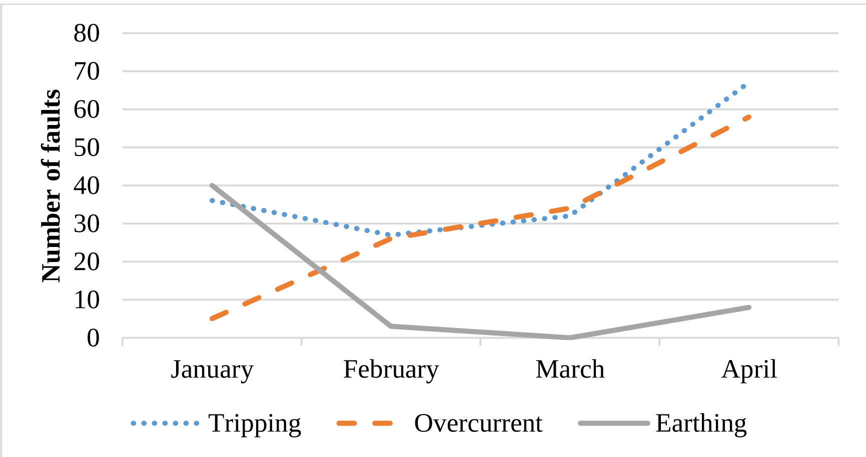

Figure 1: Sample stress faults statistics recorded from January–April 2018 at Ilala Substation (Source: TANESCO).

To maintain the operation status of SDN transformers, intelligent maintenance process is needed. Figure 2 represents procedures that are currently used by maintenance personnel to facilitate maintenance process. In the existing maintenance process, failure diagnosis and maintenance decisions are performed manually based on the expert experience or standards. Currently, most transformers installed in SDN are oil-immersed transformers; maintenance personnel compare the gas content in the oil at a particular time with the pre-defined value to diagnose its status. However, the gas content in the oil varies with age and level of oil in the transformer. Therefore, relying on the pre-defined value is not feasible. Given this observation, the system architecture for maintenance process needs to be reconfigured to include some remote sensors, controllers, and reliable communication networks to realize intelligent maintenance automation in SDN.

Maintenance strategies can be divided into two categories. The first category, which is time-based, is called preventive maintenance. In this category, maintenance of the equipment is performed based on the time interval specified regardless of the condition of the equipment. This process may introduce unnecessary costs to utility companies due to the replacement of the equipment prematurely (De Faria et al. 2015).

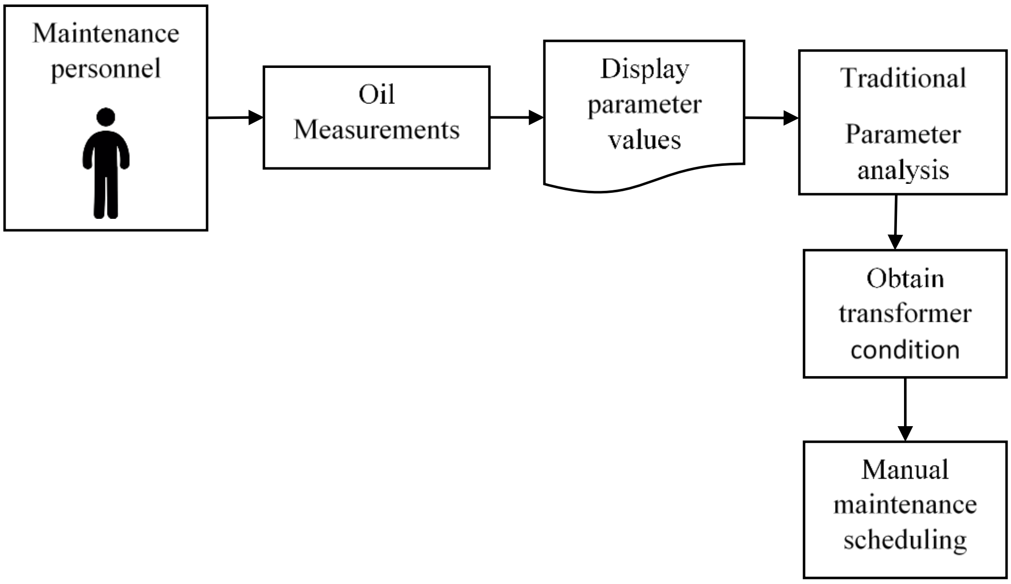

Figure 2: Current system architecture for transformer maintenance process in the Tanzania electrical secondary distribution network.

The second category, which is condition-based, is called predictive maintenance. In this category, the maintenance schedule is based on the condition of the equipment. Usually, a condition assessment of the equipment is performed before the maintenance work is performed (Noman et al. 2019). To achieve condition-based predictive maintenance in the smart grid, one should continuously monitor, analyze, and predict the status and fault of the equipment. This process allows timely maintenance action to prevent long downtime, occurrence of further damage, and unexpected failure of the equipment in the power system.

Various condition monitoring methods, such as oil analysis, thermography, vibration analysis, ultrasonic based, and electrical current monitoring, are available (Islam et al. 2017). Oil analysis through DGA (dissolved gas analysis) is widely used in assessing the condition of the oil-immersed transformer (Li and Li 2017). The typical gases that are frequently used to flag the condition of this type of transformer include hydrogen (H_2), carbon monoxide (CO), methane (CH_4), ethylene (C_2H_4), water (H_2O), acetylene (C_2H_2), and ethane (C_2H_6) (Bustamante et al. 2019). Several methods have been proposed for fault identification based on gas dissolved in oil (Gouda et al. 2018). These methods are categorized into six groups: Key gas method (KGM), Doernenburg ratio method (DRM), Rogers's ratio method (RRM), IEC ratio method (IRM), Duval triangle method (DTM), and Duval pentagon method (DPM). Table 1 summarizes the methodology and limitations of each method.

Table 1: Methods for oil-immersed transformer failure mode identification

S/N	Category	Methodology	Limitations
1.	KGM	-Based on gas percentage - Simple to use	- Identifies up to four types of fault - Low accuracy - Not reliable tool for fault analysis
2.	DRM	- Ratio-based	- Uses a few ratio test - Correctly identifies up to 3 faults
3.	RRM	- Ratio-based	-Requires a significant amount of gases - Correctly identifies up to 5 faults
4.	IRM	- Ratio-based	- More range of ratio test - Correctly identifies up to 6 fault
5.	DTM	- Ratio-based -Graphical interpretation	-Cannot be used to identify the incipient fault
6.	DPM	- Ratio-based -Graphical interpretation	- Does not specify a normal state

The methods in Table 1 are simple and easy to implement. However, evaluation of the condition of equipment at its early stage is still challenging. Lack of intelligence in traditional methods makes it challenging to predict multiple faults that may occur at the same time.

Various equipment failure prediction methods have been developed in the field of condition-based predictive maintenance. The accuracy of the maintenance automation system depends on the chosen models. Choosing the accurate model is essential since an inaccurate model may lead to the wrong decision-making process, hence imposing unnecessary costs to the utility company. Equipment fault prediction models are divided into three categories (Peng et al. 2010). The first category includes a physical model, which employs a mathematical model to realize a functional mapping between equipment failure indicators and its condition. The second category is called knowledge-based model, which transforms the expert knowledge into rules to analyze the condition of the equipment. The last category, data-driven model, depends on the data collected from sensors, and tracks the pattern that is further correlated with different equipment fault conditions. Due to the widespread deployment of low-cost sensors and the internet of things, some researchers recommend the use of data-driven methodology (Baptista et al. 2018).

The data-driven model can further be divided into two categories: statistical and artificial intelligence (AI) approaches. The former approach, which includes several statistical methods (e.g., Markov models), gives promising results for short-term prediction, but cannot guarantee the accuracy of long-term prediction (Haomin et al. 2014). The AI approach attempts to learn the equipment degradation patterns from the available observations. Recently, several AI methods have been used in predictive maintenance: artificial neural network (ANN), support vector machine (SVM), fuzzy logic, expert system, paradigm inference, and grew theory.

The ANN consists of a larger number of simple and highly interconnected processing units called neurons (Elsheikh et al. 2019). This network is capable of using sensory information to detect equipment fault and classify their functional conditions. The most important characteristic of ANN is its high speed in modeling processes and systems from the actual data. This characteristic makes ANN the ideal tool for prediction and modeling of different systems with irregular time series data (Carvalho et al. 2019). The ANN requires a large amount of dataset to train and test the system. The Tanzania utility company measures the transfomer parameter data once or twice a year, making it challenging to collect enough historical data for machine learning processing. Therefore, the data interpolation techniques become important to fully exploit the capabilities of ANNs.

Data interpolation is the process of generating new dataset from the known dataset. This process is important for the system with a limited number of historical data (Liu et al. 2019). Artificial intelligence methods for prediction require big data for training. However, acquiring enough DGA historical data may be relatively challenging. As a result, researchers have proposed several methods for data interpolation: k-Nearest Neighbours; Missforest; and multivariant imputation by chained equation methods (MICE), which this study has adopted (Cihan and Ozger 2019). Waljee et al. provided a comprehensive comparison of these methods (Waljee et al. 2013), and showed that the Missforest method is the efficient method for data interpolation. However, the practical application of Missforest for a multivariant dataset is challenging. The MICE methods, which are incorporated into the R statistical software, are powerful for data interpolation in multivariant systems. Therefore, using a suitable interpolation approach, the (interpolated) DGA data may further be processed to predict the condition status and fault type of oil-immersed transformers. But the traditional method for the DGA data processing is not feasible in a smart grid network. We need an intelligent machine learning processing to reduce human interventions in the SDN.

Significant work has been done by researchers to establish machine learning methods for transformer fault prediction and classification (Ghoneim 2018, Song et al. 2018, Zheng et al. 2018, Lin et al. 2018, Liu et al. 2019, Jiang et al. 2019 , Elânio Bezerra et al. 2020, Zeng et al. 2020). Most of the proposed artificial intelligent methods are based on single gas ratio interpretation methods (IEC 60599 2015). Consequently, a limited number of the fault types can be analyzed, hence making such methods rather weak in practical situations. Considering this limitation, we have developed a machine learning method for prediction of transformer health and for classification of transformer faults using the available DGA data. Also, based on the challenge driven education approach, expert knowledge from the utility company was acquired to improve our method. The DGA data was used to model the multiple layer artificial neural network (MLANN). Thereafter, the model was integrated with the hybrid fault classification method to maximize the number of analyzed faults of the transformer.

Materials and Methods

Case study and analysis

In the study conducted at the Tanzania electricity distribution and transmission network, the authors realized that the existing architecture in Figure 2 does not support intelligent maintenance automation. Therefore, this study proposed a system architecture to address the challenges as shown in Figure 3. The existing traditional architecture of Tanzania Electricity Supply Company Limited (TANESCO) demands manual data collection and analysis for transformer fault classification and maintenance scheduling. These procedures are time-consuming, and may contain some human-related errors. The proposed system architecture consists of automated sensors that can measure the health conditions of the transformer and send them to the intelligent embedded controller through a wireless network. In an intelligent embedded controller, the health status of the transformer was predicted and the transformer faults were determined using machine learning approaches. Thereafter, the maintenance was scheduled based on the prediction. Guided by the reliable communication architecture proposed by Chugulu and Simba (2019), the transformer processed information was then transferred to the control centre for subsequent actions. The proposed architecture opens up the possibility for the deployment of machine learning approaches for maintenance automation in the Tanzania electricity network.

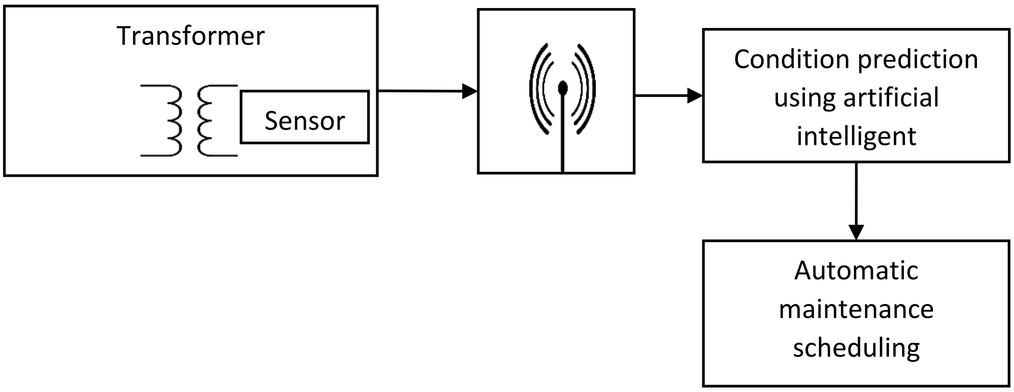

Figure 3: Proposed system architecture for transformer fault prediction in Tanzania secondary distribution network.

Sample data collection and preprocessing

The transformer historical data from 2016 to 2020 were collected from several power transformers in Tanzania as shown in Table 2. The power rating ranges from 15 kVA to 30 kVA. The authors collected 143 sets of time series historical data from the utility company. The collected data were concentration of dissolved gases in the transformer insulation oil. The concentration is measured in parts per million (ppm) using the weight of moisture divided by the weight of oil. The moisture content in oil lowers the insulating system dielectric strength and allows flashover that can damage a transformer. For mineral oil, a generally accepted maximum moisture content is 35 ppm. To acquire adequate data for prediction, the MICE computation methods were applied using the R statistical computing software version 4.0.5. The outliers were removed from the dataset as they could affect the normalization process (Pearson 2002). The decimal scaling method was used in the data normalization process (Patro and Sahu 2015).

Table 2: Sample DGA historical data in parts per million

CO_2	H_2	CO	CH_4	C_2H_6	C_2H_2	C_2H_4	H_2O
2466	9	200	16	25	0.0	32	3
2440	5	169	15	18	3.0	25	0
416	5	12	5	12	2.5	5	0
2797	5	289	4	25	1.5	6	1
4897	6	497	12	7	0.0	8	18
691	5	152	3	9	0.0	3	26
430	5	23	3	10	0.5	3	23
8333	9	1269	10	13	0.5	11	22
3368	7	141	1	44	0.0	11	22
1541	5	242	13	14	0.0	33	17
3831	15	604	6	19	3.0	9	37
1645	5	98	3	32	1.5	37	28
1280	8	221	2	9	0.5	3	17
2982	5	180	3	7	0.5	7	20
6104	8	682	10	15	0.5	1	2
4251	5	316	3	21	1.5	1	24
9405	11	819	18	18	0.0	2	22
4264	17	991	6	13	0.5	4	20
1157	5	114	6	24	0.0	5	1
6579	5	885	15	17	1.0	25	1
7240	7	1013	18	12	0.5	21	15
1098	5	55	3	11	1.0	5	23
3264	9	740	76	8	0.0	38	1
3273	8	786	10	12	0.5	4	16
2556	5	140	76	417	0.0	38	1
2495	5	159	95	481	0.0	40	19
2245	5	162	2	14	0.5	31	18
4255	9	184	9	41	0.0	4	1
2884	12	144	7	27	0.0	5	21
1157	5	114	6	24	0.0	5	1
5853	14	1034	7	14	0.5	6	22
1864	5	185	2	21	0.0	33	1
1832	5	168	2	16	1.0	32	23

Generally, the data collected from the oil-immersed transformer included H_2, CO, CH_4, C_2H_6, C_2H_2, C_2H_4, and CO_2. Guided by the expert knowledge, water (H_2O) was found to be the dorminant parameter that affects the life of the transformer in the SDN, especially during the rainy season (Bousdekis et al. 2018). Also, an observation was made that the transformer material may add water to the oil, hence, promoting premature transformer cellulose ageing. Noting these observations from the expert knowledge, the authors used the eight transformer parameters for condition status prediction and fault analysis.

Prediction model selection

Several prediction models have been proposed in the literature (Li and Li 2017, Bousdekis et al. 2018). To investigate the accuracy of the prediction models based on the available dataset, the authors tested the SVM, KSVM, and MLANN. The implementation steps (Figure 4) include DGA data collection from TANESCO, data analysis, and model identification. The transformer fault prediction was performed on each model by using the same dataset, and the performance was evaluated using mean absolute error (MAE), mean absolute parcentage error (MAPE), and root mean square error (RMSE). In this study, the R-studio software was used in the model training and testing processes.

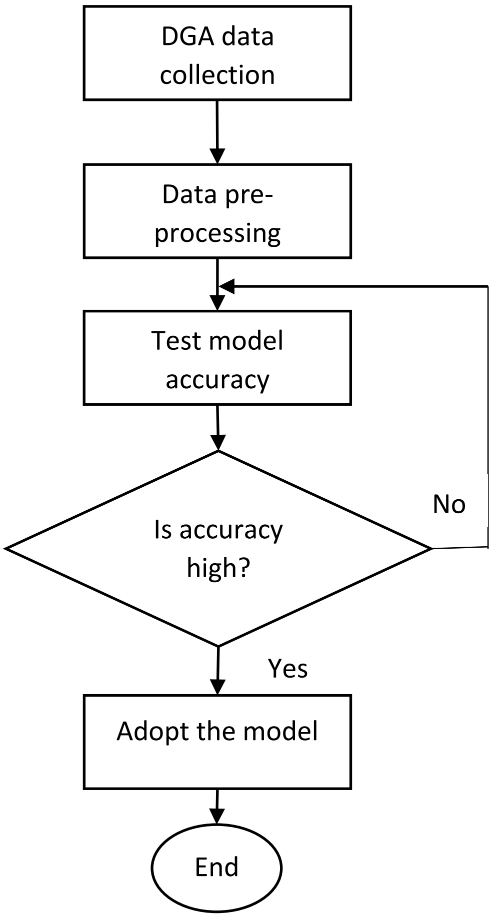

Figure 4: Steps for selection of prediction model.

The accuracy of the models SVM, KSVM, and MLANN identified by subtracting MAPE from 100% were found to be 96.72%, 97.50%, and 97.53%, respectively. Therefore, MLANN was selected for its high accuracy compared with the other models.

Transformer condition prediction based on MLANN

One thousand DGA sample data were divided into two parts, 800 training samples and 200 testing samples. The MLANN structure consists of the input layer, hidden layer, and output layer. In this study, the input layer of MLANN contained ten input parameters, four hidden layers, one output layer, and a logistic activation function as shown in Figure 5. The output of the model provides information regarding the transformer status. The authors divided the transformer conditions into four types, namely normal, warning, caution, and critical. In case of the abnormal condition, further analysis was performed to determine the fault type based on the characteristics of the input parameters.

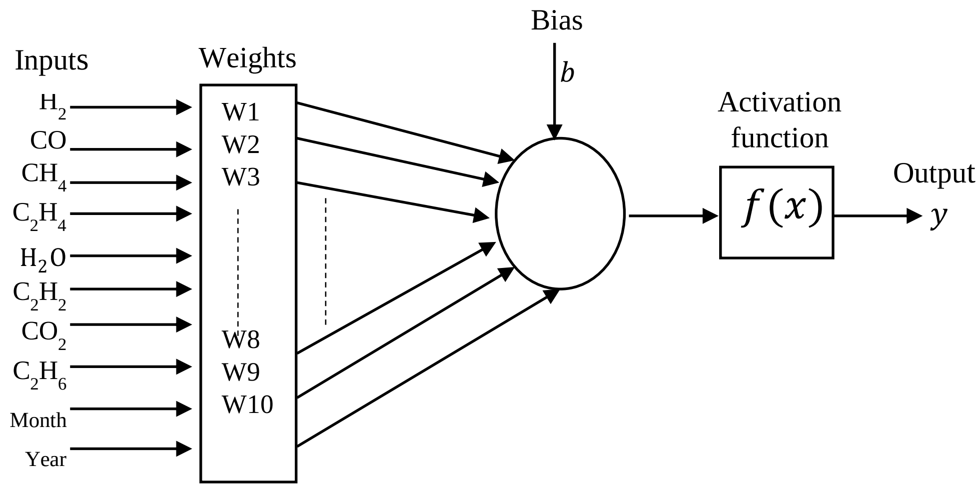

Figure 5: Structure of MLANN state prediction based on DGA data.

Transformer fault analysis based on hybrid method

According to the IEC 60599 standards, transformer fault types are classified into six groups: Partial discharge, low-energy discharge, high-energy discharge, low-temperature thermal fault, medium-temperature thermal fault, and high-temperature thermal fault (IEC 60599, 2015). Scholars have proposed several oil-immersed transformer fault identification methods (Gouda et al. 2018). Despite the attempts, the proposed methods disregard the effect of water content. Therefore, in this study, expert knowledge was introduced to address the challenge. This approach encourages the application of real practical analysis of data to formulate testable logic for fault identification in oil-immersed transformers. Expert knowledge can be acquired through challenge-based education that the current study has adopted (Kalinga et al. 2017).

In this work, the authors have proposed the hybrid method that integrates the IEC 60599 method, gas ratio method, and expert knowledge-based method for the transformer fault identification. The input for the fault identification is shown in Table 3.

Table 3: Transformer fault analysis input parameters

Fault identification method	Input parameters
IEC ratio	C_2H_2/ C_2H_4, CH_4/ H_2, C_2H_4/ C_2H_6
Gas ratios	CO_2/ CO, C_2H_2/ H_2
Expert knowledge	CO, CO_2, H_2O

The process of identifying the generalized fault identification methods includes five steps: data collection, data processing (analysis of the available faults and records), reviewing of the existing methods for transformer faults identification, test of each method with available DGA data, and analysis of the results as shown in Figure 6.

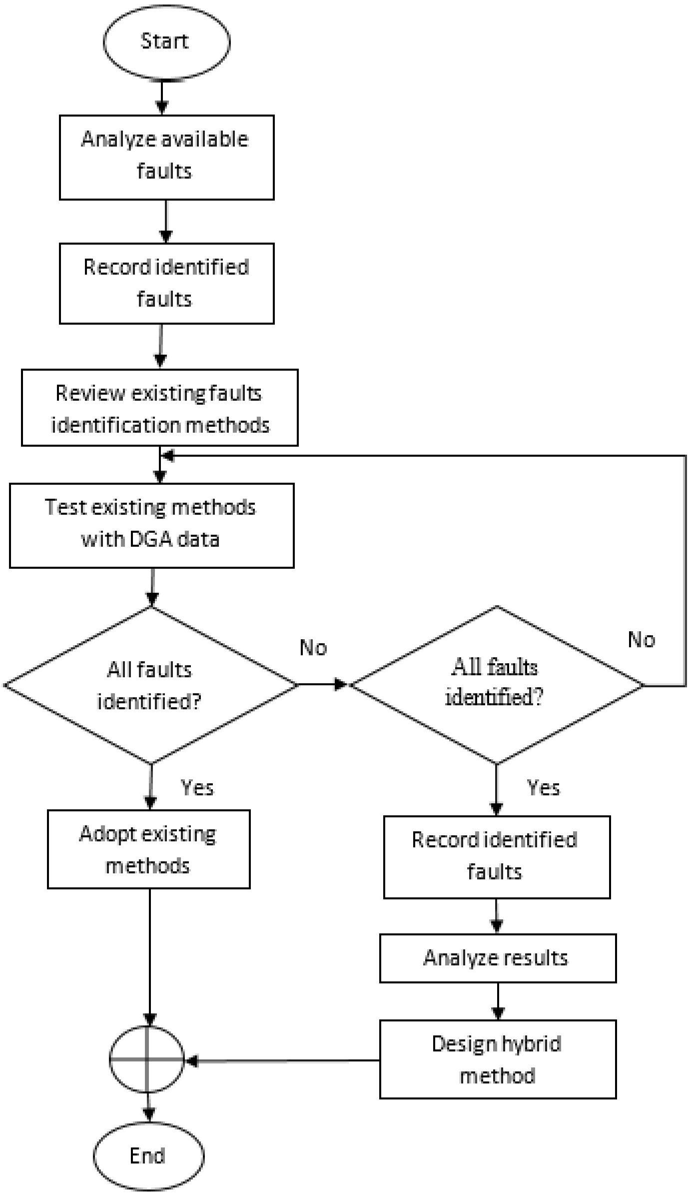

Figure 6: Fault identification method selection steps based on DGA.

Results and Discussion

Fault identification

The R-studio software was used to simulate faults identification of the existing and proposed hybrid method. The DGA data from TANESCO was used for training and testing the performance of the existing and proposed hybrid method. Table 4 presents results of ten types of faults tested on the existing and proposed hybrid method. The results in Table 4 show that the proposed hybrid method can identify ten faults, accounting for 100% of the tested faults. The IEC 60599 standard method identifies six faults which is 60% of the tested faults. The gas ratio and limit-based methods identify only two faults which is 20% of all tested faults. Therefore, the proposed method is more generalized and can be used in the maintenance decision-making process. These observations suggest that the proposed system may be integrated into a maintenance scheduling platform to reduce unplanned maintenance outages and human maintenance-related errors

Faults identification patterns were also simulated for each of the selected methods. The intention was to identify the most dominant faults in the recorded data from 2016 to 2020. Figure 7 shows that IEC 60599 method detects T1 and T2 as the most recurring faults in all five years. However, the T1 fault decreased from 140 times in 2019 to 60 times in 2020. Figure 8 shows that Gas Ratio method detects CDA as the most recurring fault in all five years. However, the CDA fault decreased from 180 times in 2019 to 80 times in 2020. Figure 9 shows that Limit Based method detects CIT as the most recurring fault in all five years. However, the CIT fault decreased from 210 times in 2019 to 80 times in 2020. Figure 10 shows that proposed hybrid method detects T1 as the most recurring faults in all five years. However, the T1 fault decreased from 140 times in 2019 to 50 times in 2020. These simulation results suggest that each method can identify transformer abnormalities, hence, suggesting inaccurate maintenance decision-making processes.

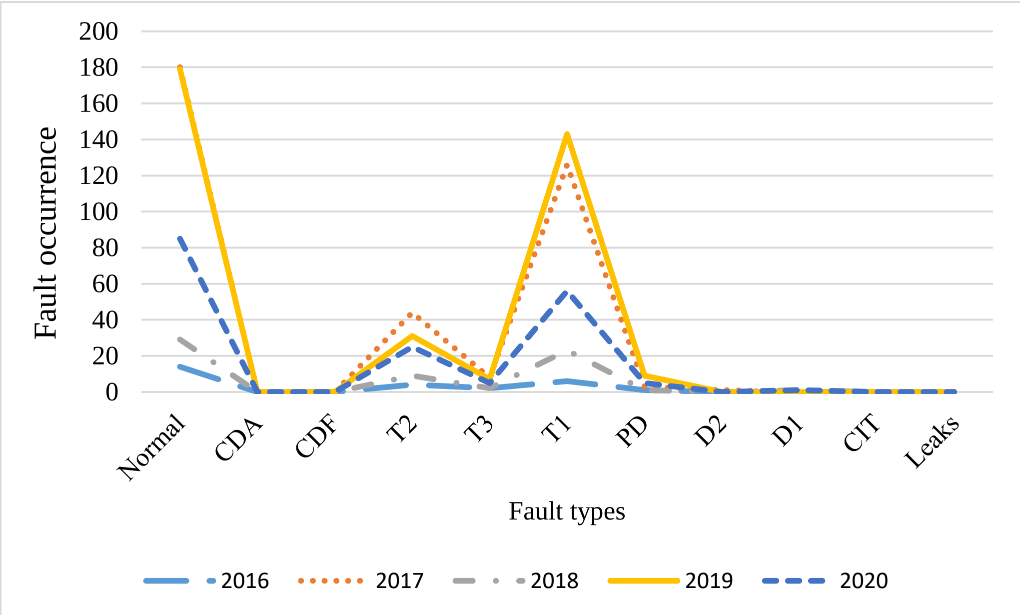

Figure 7: Faults identified by IEC 60599 method.

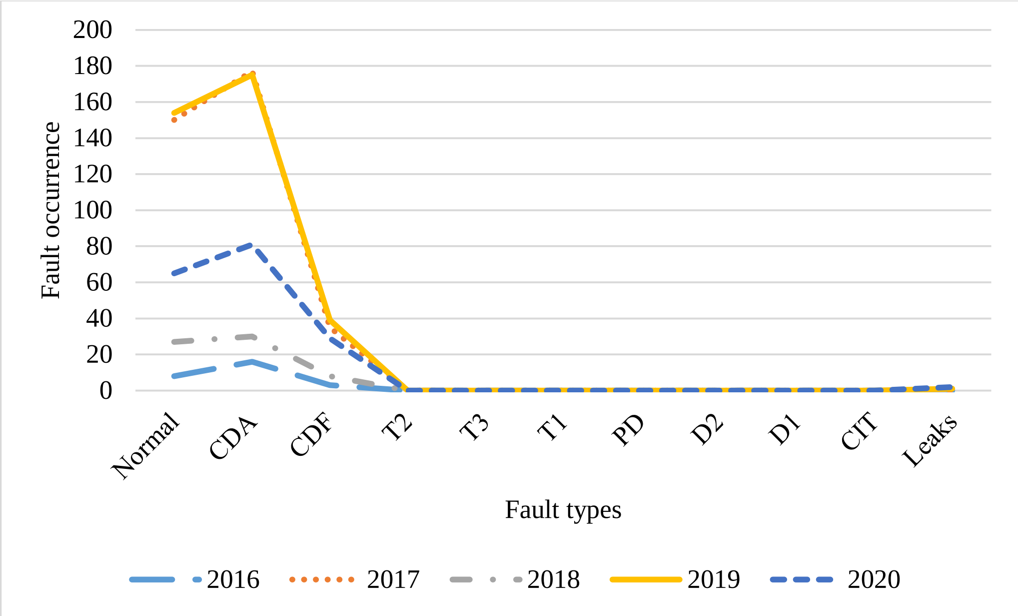

Figure 8: Faults identified by Gas Ratio method.

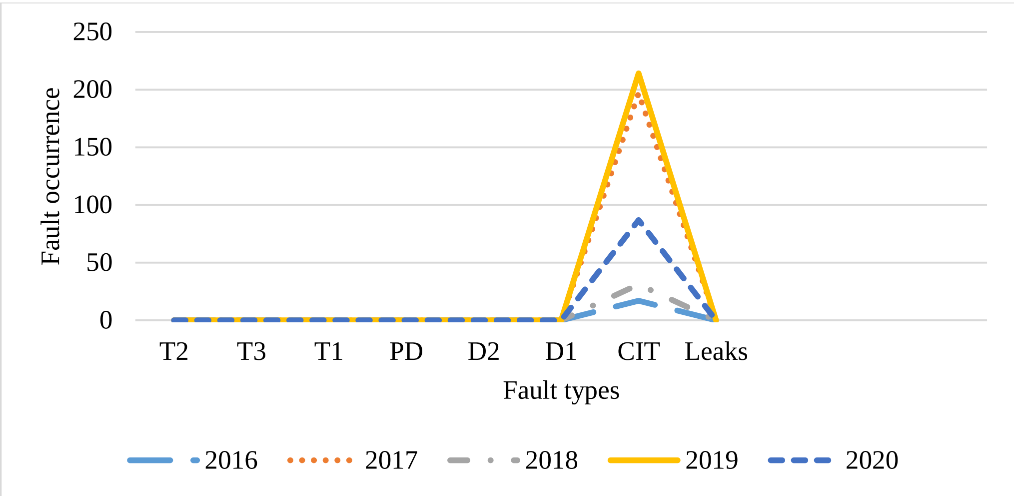

Figure 9: Faults identified by Limit-Based method.

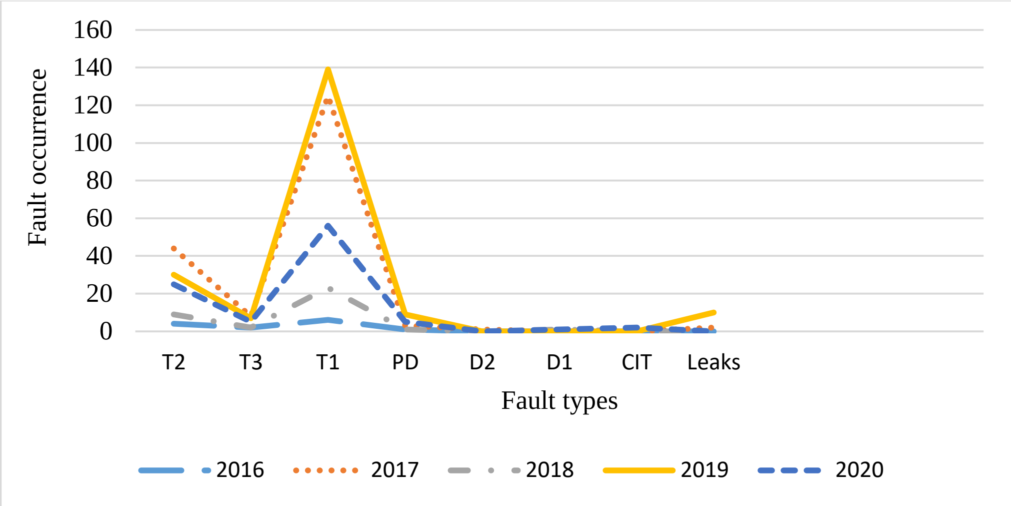

Figure 10: Fault identification based on proposed hybrid method.

Table 4: Faults Identification

No	Fault type	Description	IEC method	Gas ratio method	Limit-based method	Proposed hybrid method
1	PD	Partial discharge	√	×	×	√
2	D1	Low energy discharge	√	×	×	√
3	D2	High- energy discharge	√	×	×	√
4	T1	Thermal fault T > 300 ℃	√	×	×	√
5	T2	Thermal T > 300 ℃ – 700 ℃	√	×	×	√
6	T3	Thermal fault T > 700 ℃	√	×	×	√
7	CDA	Paper deterioration due to ageing	×	√	×	√
8	CDF	Cellulose deterioration due to other factors (stress faults)	×	√	×	√
9	CIT	Contamination in tank	×	×	√	√
10	Leaks	Water leaks into the oil	×	×	√	√

√ – Identified PD: Partial discharge D2: High-energy discharge
× – Not-identified D1: Low-energy discharge T1: Low-temperature thermal fault
T2: Medium-temperature thermal fault T3: High-temperature thermal fault

Transformer condition prediction based on MLANN

This study proposed a condition and fault prediction method by analyzing the main failure characteristics parameters of oil-immersed transformers. Based on the available data, three prediction models SVM, KSVM and MLANN, were tested to determine their performance in terms of MAE, MAPE and RMSE as shown in Table 5. Results showed that MLANN had RMSE of 0.00094 compared to 0.00101 and 0.00265 of KSVM and SVM, respectiviely. This observation suggests that MLANN has higher accuracy than KSVM and SVM models. Therefore, in transformer fault predictions, the MLANN model could give the most accurate results.

Table 5: Performance analysis of prediction models based on DGA data

Model	MAE	MAPE	RMSE
SVM	0.03270	3.2734	0.00265
KSVM	0.02493	2.4937	0.00101
MLANN	0.02493	2.4685	0.00094

Conclusion

This study focused on the maintenance automation in distribution and transmission networks in Tanzania power systems, where intelligent equipment fault identification has not previously been deployed. The authors have proposed a system architecture that allows maintenance automation in the Tanzania secondary distribution networks. Based on the DGA data, the current study reveals that MLANN is the most suitable model for transformer state prediction. Also, using the same data, the most popular transformer fault identification methods were tested and compared. Results suggest that the existing methods have a limitation of not being able to identify all type of faults. Therefore, the authors proposed the MLANN-hybrid method for prediction of transformer conditions and fault in the secondary distribution networks. The proposed method can be used in maintenance decision-making as it includes most of the identified transformer abnormalities based on oil analysis. As a possible future research avenue, there seems to be a need to design a maintenance scheduling method that integrates the transformer state prediction and hybrid fault identification methods. The maintenance decision based on the prediction can significantly reduce the costs of unexpected downtimes.

Acknowledgement

This research was conducted under the iGrid project of the University of Dar as Salaam (UDSM). The project is sponsored by the Swedish International Development Agent (SIDA). Special thanks go to TANESCO for the great collaboration it provided during the entire research period.

References

Asadzadeh SM and Azadeh A 2014 An integrated systemic model for optimization of condition-based maintenance with human error. *Reliab. Eng. Syst. Saf.* 124: 117–131.

Baptista M, Sankararaman S, Medeiros IP De, Nascimento C, Prendinger H, and Henriquesa EMP 2018 Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling. *Comput. Ind. Engin.* 115: 41-53.

Bhargava C, Sharma PK, Senthilkumar M, Padmanaban S, Ramachandaramurthy VK, Leonowicz Z, Blaabjerg F, and Mitolo M 2020 Review of health prognostics and condition monitoring of electronic components. *IEEE Access*. 8: 75163–75183.

Bousdekis A, Magoutas B, Apostolou D, and Mentzas G 2018 Review, analysis and synthesis of prognostic-based decision support methods for condition based maintenance. *J. Intell. Manuf.* 29(6): 1303–1316.

Bustamante S, Manana M, Arroyo A, Castro P, Laso A, and Martinez R 2019 Dissolved gas analysis equipment for online monitoring of transformer oil: a review. *Sensors* 19(19): 4057.

Carvalho TP, Soares FAAMN, Vita R, Francisco R da P, Basto JP, and Alcalá SGS 2019 A systematic literature review of machine learning methods applied to predictive maintenance. *Comput. Ind. Eng.* 137: 106024.

Chugulu G and Simba F 2019 Communication Architecture for automatic faults detection and clearance in secondary distribution power grid: The case of TANESCO. In *Proceedings of 2019 the 7th International Conference on Smart Energy Grid Engineering, SEGE 2019* 223–229. Institute of Electrical and Electronics Engineers Inc.

Cihan P and Ozger ZB 2019 A new heuristic approach for treating missing value: ABCIMP. *Elektron. Elektrotech* 25(6): 48–54.

Colak I, Sagiroglu S, Fulli G and Yesilbudak M 2016 A survey on the critical issues in smart grid technologies. *Renew. Sustain. Energy Rev*. 54: 396–405.

De Faria H, Costa JGS, and Olivas JLM 2015 A review of monitoring methods for predictive maintenance of electric power transformers based on dissolved gas analysis. *Renew. Sustain. Energy Rev.* 46: 201-209.

Dileep G 2020 A survey on smart grid technologies and applications. *Renew. Energy* 146: 2589–2625.

Elânio Bezerra F, Zemuner Garcia FA, Ikuyo Nabeta S, Martha de Souza GF, Chabu IE, Santos JC, Junior SN, and Pereira FH 2020 Wavelet-Like transform to optimize the order of an autoregressive neural network model to predict the dissolved gas concentration in power transformer oil from sensor data. *Sensors* 20(9): 2730.

Elsheikh AH, Sharshir SW, Abd M, Kabeel AE, Guilan W and Haiou Z 2019 Modeling of solar energy systems using arti fi cial neural network : A comprehensive review. *Solar Energy* 180: 622–639.

Ghoneim SSM 2018 Intelligent prediction of transformer faults and severities based on dissolved gas analysis integrated with thermodynamics theory. *IET Sci., Meas. Technol.* 12(3): 388–394.

Gouda OE, El-Hoshy SH, and El-Tamaly HH 2018 Proposed heptagon graph for DGA interpretation of oil transformers. *IET Gener. Transm. Distrib*. 12(2): 490–498.

Haomin C, Peng L, Xiaobin G, Aidong X, Bo C, Wei X, and Liqiang Z 2014 Fault prediction for power system based on multidimensional time series correlation analysis. *China International Conference on Electricity Distribution, CICED*. 2014-Decem(Ciced): 1294–1299.

IEC 60599 2015 Mineral oil-impregnated electrical equipment in service–Guide to the interpretation of dissolved and free gases analysis 3: 1–69.

Islam M, Lee G, and Hettiwatte SN 2017 A review of condition monitoring techniques and diagnostic tests for lifetime estimation of power transformers. *Electr.l Eng.* 100(2): 581-605.

Jiang J, Chen R, Chen M, Wang W, and Zhang C 2019 Dynamic fault prediction of power transformers based on hidden Markov model of dissolved gases analysis. *IEEE Trans. Power Delivery* 34(4): 1393-1400.

Kalinga E, Ibwe K and Mvungi N 2017 The role of challenge driven learning in activating industry-academia cooperation in low-income countries. In *ICERI 2017 Proceedings* 8158–8166.

Li S and Li J 2017 Condition monitoring and diagnosis of power equipment : review and prospective. *High Voltage* 2: 82–91.

Lin J, Su L, Yan Y, Sheng G, Xie D and Jiang X 2018 Prediction method for power transformer running state based on LSTM_DBN Network. *Energies*. 11(7): 1880.

Liu J, Zheng H, Zhang Y, Li X, Fang J, Liu Y, Liao C, Li Y and Zhao J 2019 Dissolved gases forecasting based on wavelet least squares support vector regression and imperialist competition algorithm for assessing incipient faults of transformer polymer insulation. *Polymers* 11(1): 85.

Mnyanghwalo D, Kawambwa S, Mwifunyi R, Gilbert GM, Makota D and Mvungi N 2019 Fault detection and monitoring in secondary electric distribution network based on distributed processing. In *2018 20th International Middle East Power Systems Conference, MEPCON 2018-Proceedings* 84–89. Institute of Electrical and Electronics Engineers Inc.

Noman MA, Nasr ESA, Al-Shayea A, and Kaid H 2019 Overview of predictive condition based maintenance research using bibliometric indicators. *J. King Saud Univ. Eng. Sci.* 31(4): 355-367.

Patro SGK and Sahu KK 2015 Normalization: a preprocessing stage. *IARJSET*. 20–22.

Pearson RK 2002 Outliers in process modeling and identification. *IEEE Trans. Control Syst. Technol.*. 10(1): 55–63.

Peng Y, Dong M, and Zuo MJ 2010 Current status of machine prognostics in condition-based maintenance : a review. *Int. J. Adv. Manuf. Technol.* 50*:* 297–313.

Song H, Dai J, Luo L, Sheng G, and Jiang X 2018 Power transformer operating state prediction method based on an LSTM Network. *Energies* 11(4): 914.

Waljee AK, Mukherjee A, Singal AG, Zhang Y, Warren J, Balis U, Marrero J, Zhu J, and Higgins PDR 2013 Comparison of imputation methods for missing laboratory data in medicine. *BMJ Open*. 3(8): e002847.

Zeng B, Guo J, Zhang F, Zhu W, Xiao Z, Huang S, and Fan P 2020 Prediction model for dissolved gas concentration in transformer oil based on modified grey wolf optimizer and lssvm with grey relational analysis and empirical mode decomposition. *Energies*. 13(2): 422.

Zheng H, Zhang Y, Liu J, Wei H, Zhao J, and Liao R 2018 A novel model based on wavelet LS-SVM integrated improved PSO algorithm for forecasting of dissolved gas contents in power transformers. *Electr. Power Syst. Res.* 155: 196–205.