IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 24, 2018, accepted February 14, 2018, date of publication February 27, 2018, date of current version March 15, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2809436

Transformer Fault Condition Prognosis Using Vibration Signals Over Cloud Environment

MEHDI BAGHERI, (Member, IEEE), AMIN ZOLLANVARI, (Member, IEEE), AND SVYATOSLAV NEZHIVENKO

Department of Electrical and Computer Engineering, School of Engineering, Nazarbayev University, Astana 01000, Kazakhstan

Corresponding author: Mehdi Bagheri (mehdi.bagheri@nu.edu.kz)

ABSTRACT On-line monitoring and diagnosis of transformers have been investigated and discussed significantly in the last few decades. Vibration method is considered as one of the non-destructive and economical methods to explore transformer operating condition and evaluate transformer mechanical integrity and performance. However, transformer vibration and its evaluation criteria in transformer faulty condition are quite challenging and are not yet agreed upon. At the same time, with the advent of IoT facilities and services, it is expected that classical diagnosis techniques will be replaced with more powerful data-driven prognosis methods that can be used efficiently and effectively in smart monitoring. In this paper, we first discuss in detail an analytical approach to the transformer vibration modeling. Nevertheless, precise interpretation of transformer vibration signal through analytical models becomes unrealistic as higher harmonics are mixed with fundamental harmonics in vibration spectra. Therefore, as the next step, we aim to support the Industry 4.0 concept by utilizing the state-of-the-art machine learning and signal processing techniques to develop prognosis models of transformer operating condition based on vibration signals. Transformer turn-to-turn insulation deterioration and short circuit analysis as one the most important concerns in transformer operation is practically emulated and examined. Along with transformer short-circuit study, transformer over and under excitations are also studied and evaluated. Our constructed predictive models are able to detect transformer short-circuit fault in early stages using vibration signals before transformer catastrophic failure. Real-time information is transferred to the cloud system and results become accessible over any portable device.

INDEX TERMS IoT in power system, online transformer assessment, prognosis, vibration analysis, signal modeling, prediction, regression.

I. INTRODUCTION

Asset management is more important now than ever for industries. Indeed utility management and all industrial and system engineers are more emphasizing these days on intelligent predictive assessment rather than asset restoration after catastrophic failure. For electrical items, *smart predictive assessment* becomes meaningful when online asset prognosis techniques are taken into consideration. At the same time, emerging technologies such as Internet of Things (IoT) and cloud computing have empowered various smart monitoring systems. With development of technology, specifically in the context of smart grid, utilizing IoT in different aspects of power system is becoming popular.

Equipment are expected to work with IoT protocols, transfer, restore and analyze data through cloud environment; however, making automated algorithms, managing and connecting equipment to the cloud system and real-time evaluation of their performance via cloud are still significant challenges ahead. Furthermore, it is well-known that power transformer is one of the most expensive equipment among all electrical items. This valuable equipment is in service in various climates as well as different electrical and mechanical conditions [1]. Based on this fact, transformers are continually facing enormous hazards over the course of operation. Yielding information continuously from transformer condition and having a reasonable understanding about internal mechanical stability is vitally important for the system operators.

Different on-line and off-line methods have been introduced and implemented for transformer mechanical integrity assessment [2]–[7]. However, practical key methods such as Frequency Response Analysis (FRA) are still

performing off-line. Implementation of on-line FRA is seriously under study and development. It will bring huge benefits to industry if it is performed perfectly. Meanwhile, FRA is not the only method providing information on transformer mechanical integrity. Online short circuit impedance measurement [8], transformer sound analysis (specifically ultrasound interpretation) [9], communication based methods as well as a technique based on locus diagram have been introduced for transformer winding deformation detection [10]–[13]. Transformer vibration assessment is also a well-known technique to evaluate transformer mechanical integrity.

Transformer vibration can be considered to be the repetitive movement of transformer inner parts that are covered by the transformer tank, or perhaps the movement of active part in dry-type transformer. Such movements occur around a reference position. The reference position is where the transformer attains once it is out of service. This method is considered as the most economical real-time method to analyze transformer mechanical integrity as it does not need any complicated test setup to be deployed for transformer monitoring. Nevertheless, the predictive performance of vibration-based techniques and their capacity and benefits for transformer evaluations have not been as widely discussed in the literature as other methods. This state of affairs can be attributed, in part, to the lack of a unique interpretation method for vibration data and, in part, to the challenging habit of noisy vibration signal and its dependency on various factors.

Vibration might be interpreted by using parameters such as winding displacement, velocity, and acceleration. Although it is more common to choose acceleration as the parameter of interest for vibration analysis, no objective criteria has been introduced for this technique. In fact, vibration signals are fully mixed with undesirable environmental, background, or even auxiliary equipment oscillations. At the same time, unpredictability of vibration signal noise in different transformers often prohibits the development of a unique reliable interpretation system. For instance, in large power transformers, mixing the active part vibration of transformer with oil pump, tap-changer, fans, cooling system, makes the transformer vibration signal interpretation challenging and unreliable. This scenario becomes more complicated when a fault is initiated in transformer. The cumulative and complex effect of these factors on vibration signal renders analytical methods complicated and often intractable. On the other hand, machine learning and signal processing techniques lend themselves to finding data-driven solutions in similar complicated scenarios in various fields of study. These data-driven methods are primarily used in this study to distinguish fault initiation and perform prognosis.

Different studies have been conducted to study transformer vibration characteristics. Garcia *et al.* [15] has studied vibration analysis of transformer tank and its benefits to recognize transformer winding deformation. Another worthy study by Ji *et al.* [16] introduces an online technique called On-load Current Method (OLCM) to distinguish the transformer core through winding vibration when the transformer is in service. In our previous study [17], the vibration analysis is presented and discussed as an on-line method in transformer winding deformation recognition. In a study by Saponara *et al.* [18], predictive diagnosis of high-power transformer via networking vibration measuring nodes is discussed. Vibration technique is also used in [19] for predictive diagnosis of uninterruptible power supplies in safety-critical applications. Both [18] and [19] discuss industrial applications of vibration analysis method and its practical performance.

In [20] three different indices are introduced and discussed to analyze transformer mechanical assessment. Vibration correlation to find transformer winding conditions is also discussed in [21]. Despite the remarkable progress made by these studies, no robust and prognosis method for analyzing transformer vibration measurements has been developed. Moreover, it is almost impossible to load and adopt the methods developed in above studies with IoT and cloud computing.

In this study, we first discuss in detail an analytical approach to the transformer vibration modeling. We then use the state-of-the-art machine learning and signal processing techniques to construct remarkably accurate predictive vibration models of transformer conditions. Using analysis of vibration signals, we study two problems: (1) the transformer under and over excitations; and (2) inter-turn short-circuit. In the first case, observations over the entire range of interest for the injected voltage (320-440V) are available in the model training stage. A predictive model is then developed and used to predict excitation voltage in a similar range of injected voltage. In the second case, the training stage contains only observations collected from load currents ranging from 3-12A. However, we intend to apply the constructed model to predict short-circuit currents of 13 and 15A. In both cases, we model vibration signals by a mathematical form, which is the sum of sinusoids (of order M) with unknown amplitudes, frequencies, and phases. A nonlinear least square technique is then used to estimate the unknown parameters of this mathematical form [23]. However, the difference between the predictive objectives of the constructed models in these two cases leads to different model-feature selection process used in the training stage (and quite different structure of the final constructed models).

In the first case study, we use a nested cross-validation for model selection (i.e., to determine M and choosing linear or nonlinear model). By applying cross-validation external to the feature selection process, we prevent the selection bias in the feature selection stage, which potentially leads to more accurate predictive models [28]. Once the predictive model is constructed, the model is validated on an independent set of observations collected from a similar range of excitation voltage used in the training stage.

In the second case study, the constructed model will be used in predicting short-circuit currents that have not been available during the training phase (13 and 15A). Therefore, we divide the training data to a range of 3-10A and optimize

the performance of the constructed models on the observations collected for 11 and 12A. This practice has simulated the real scenario where independent set of observations associated with a range of load currents that are unseen during the training phase would need to be predicted. After model selection phase, we use the full training data (3-12A) to develop the final model and validate its predictive capacity in a range of short-circuit currents (13 and 15A). All methods are implemented over a cloud system and able to evaluate the transformer condition simply via IoT protocols. Evaluated results are then provided on different protocoled devices for prognosis of transformer active parts.

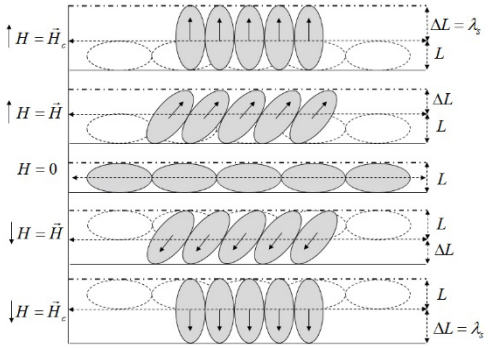

FIGURE 1. Magnetostriction of ferromagnetic material.

II. TRANSFORMER CORE VIBRATION MODELING

Magnetostriction is an expression for the length alteration in a ferromagnetic material once it is under magnetization. The direction and extension of magnetization will impose the stress over the ferromagnetic material. In other words, in the absence of the magnetic field, the magnetic material is in relaxation mode and its magnetic domains are randomly oriented. Activating the magnetic field will cause that the magnetic domains are starting to align themselves with the imposed field and get orientation (see Fig. 1). The anisotropy principle axes would be then collinear to the magnetic field. A large mechanical deformation in microscopic scale is obtained due to increasing of the magnetic field over a magnetic material. This deformation is reached to its maximum value once the magnetic domains are reached to their saturation points. This was explored by J. Jole in 1842. Undoubtedly, magnetostriction is very much effective and considerable for ferromagnetic and ferrimagnet materials rather than paramagnets and diamagnetic. Transformer core vibration is mainly due to the magnetostriction. Suppose that a ferromagnetic core sheet with thickness of b, width of w, and length of L is exposed to the magnetic field [16],

$$U_0 \sin \omega t = -N_w \frac{d\phi}{dt} = -N_w A_c \frac{dB}{dt}, \tag{1}$$

where, U_0 is the voltage source applied to the winding, ω is the angular frequency, N_w is the number of winding turns, B is the magnetic induction, and A_c is the cross section area of single core laminate. Therefore, the magnetic induction is calculated as [16]

$$B = \frac{-U_0}{N_w A_c} \int \sin \omega t dt = \frac{U_0}{N_w A_c \omega} \cos \omega t = B_0 \cos \omega t, \tag{2}$$

where B_0 is the maximum induction magnitude. B_0 is less or equal to the induction saturation level (B_s) [16]

$$B_0 = \frac{U_0}{N_w A_c \omega} \leq B_s. \tag{3}$$

In addition, the field density and magnetic induction is related through the magnetic permeability as

$$B = \mu H, \tag{4}$$

where μ is the magnetic permeability, and H is the magnetic intensity. H_c is the saturated magnetic intensity and is obtained where H obtains its maximum value:

$$B_s = \mu H_c. \tag{5}$$

Thus, the applied magnetic field intensity and the saturated induction magnetic intensity can be given by

$$B = \frac{B_s}{H_c} H. \tag{6}$$

The applied magnetic field intensity is obtained by replacing (2) in (6); to wit,

$$H = \frac{H_c B_0}{B_s} \cos \omega t. \tag{7}$$

Any alteration in the core laminate length is initiated by the changes in the field magnetic intensity. Thus, the core laminate maximum movement due to the field intensity changes is given by [16]

$$\begin{aligned} x_{core} &= \frac{dL}{L} = \frac{\lambda_s}{H_c^2} \int_{-H}^{H} H dH = \frac{2\lambda_s}{H_c^2} \int_{0}^{H} |H| dH = \frac{\lambda_s H^2}{H_c^2} \\ &= \frac{\lambda_s}{H_c^2} \frac{H_c^2}{B_s^2} B_0^2 \cos^2 \omega t = \frac{\lambda_s}{H_c^2} \frac{H_c^2}{B_s^2} \left(\frac{U_0}{N_w A_c \omega} \right)^2 \cos^2 \omega t \\ &= \frac{\lambda_s U_0^2}{B_s^2 N_w^2 A_c^2 \omega^2} \cos^2 \omega t, \end{aligned} \tag{8}$$

where λ_s is the maximum magnetostriction value. Therefore, the core laminate acceleration is obtained as

$$\ddot{x}_{core} = \frac{d^2 x_{core}}{dt^2} = -\frac{2\lambda_s L U_0^2}{B_s^2 N_w^2 A_c^2} \cos 2\omega t. \tag{9}$$

Equation (9) shows that the magnitude of the transformer core vibration is proportional to the excitation voltage square. Also, the frequency of core vibration is matched with the second harmonic order of the voltage fundamental frequency. Hence, for the systems with the fundamental frequency of 50 Hz, the transformer core laminates vibrate with 100 Hz as the fundamental transformer core vibrational frequency, and for the systems with 60 Hz, they vibrate with 120 Hz. Furthermore, higher vibrational harmonic orders would be created in transformer core due to the collision of the core

laminates. In fact, the extensions and contractions of the core laminates will cause that they hit each other regularly. The vibration frequencies of this clash are higher than the fundamental vibrational frequency. The transformer core will experience 200, 400, and 600 Hz or even rarely higher harmonic orders in its vibrational spectrum due to the core laminates collision. The magnitudes of the higher harmonic orders in transformer core vibration are quite random; that is to say, they are not necessarily proportional to the excitation voltage square value.

Natural mechanical resonance frequencies as well as technical assembly considerations, Lorentz forces and also magnetic core quality can easily influence entire core vibration spectrum. Hence, a range of frequencies are feasible to be explored in the transformer core vibration spectrum; however, second harmonic 2ω, is considered as the fundamental transformer core vibrational frequency.

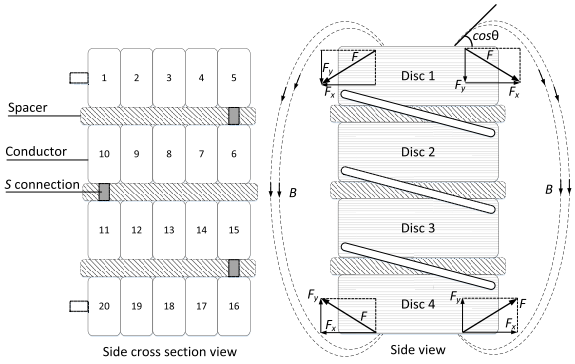

FIGURE 2. Transformer winding schematic, side cross section view, electromagnetic force direction, and flux lines.

III. TRANSFORMER WINDING VIBRATION MODELING

Summation of linkage and leakage fluxes are making the entire flux in transformer active part. Technically, a major part of linkage flux and leakage flux flows through the transformer core and the transformer winding and active part gaps, respectively. In other words, vibration in the transformer core and winding is mainly initiated by linkage and leakage flux, respectively. Indeed, transformer winding vibration is due to the electromagnetic force created by the leakage magnetic flux, $B_{leakage}$, flowing through the winding and the winding current. Figure 2 shows that the leakage magnetic flux direction is changed by moving from the winding sides towards winding center. Therefore, the electromagnetic force direction can be changed, and consequently mechanical force direction is changed. Moreover mechanically, the transformer winding can be modeled by spring. Hence, the spring-force model can be used to analyze the transformer winding vibration. Thus, four different scenarios can be assumed.

The first scenario stands for an impulse force over the winding (spring) without damping factor. This force will be removed during winding vibration. This is in turn a basic vibration model. The second scenario occurs when the force will remain over the transformer winding; however, we still do not consider a damping factor. Third scenario is similar to the first scenario; however, damping factor is taken into account, and fourth scenario is a complete model of transformer winding condition where the force will remain over the winding and damping factor is also taken into account for modeling. All the scenarios are discussed in this section.

A. FREE VIBRATION WITHOUT DAMPING FACTOR

In the absence of the mechanical constraints, the transformer winding can move vertically due to any impulse force strike. This is a freely movement and the transformer winding will oscillate naturally. The relation between the imposed force and the movement is represented by the unit extension of spring factor, k, and equation is obtained as,

$$F = -kx, \tag{10}$$

where F represents the force, and x is the displacement. If the spring weight is taken into account and a single impulse is imposed to the spring, according to the Newton's second law, (10) will be developed and force equation is obtained as

$$ma = F' = W - F = W - (W + kx), \tag{11}$$

where a is the acceleration factor or second derivative of displacement, and W is the winding weight. Equation (11) can be then re-written as

$$ma = \frac{W}{g}a = \frac{W}{g}\ddot{x} = W - (W + kx), \tag{12}$$

and g denotes the gravity acceleration. Rearranging (12), the differential equation for natural motion of winding is obtained as

$$\frac{W}{g}\ddot{x} + kx = 0, \quad \ddot{x} + \frac{g}{W}kx = 0. \tag{13}$$

Simplification of (13) leads to

$$\frac{gk}{W} = \alpha^2, \quad \ddot{x} + \alpha^2 x = 0. \tag{14}$$

The solution of differential equation, (14), is given as

$$x_1 = C_1 \cos \alpha t + C_2 \sin \alpha t, \tag{15}$$

where, C_1 and C_2 are the constants determining by initial condition, and α is given by

$$\alpha = \frac{2\pi}{\tau_n}, \quad \tau_n = 2\pi\sqrt{\frac{\delta_{st}}{g}}, \quad f_n = \frac{1}{2\pi}\sqrt{\frac{g}{\delta_{st}}}, \tag{16}$$

where f_n denotes the natural oscillation frequency of spring when the damping factor is ignored.

B. FREE VIBRATION WITH DAMPING FACTOR

A prevention in natural oscillation of winding is obtained when the damping factor is considered for the winding motion due to the impulsive force (instantaneous force).

Damping factor potentially prevents the winding fast oscillation. In this case, winding motion equation is given by

$$\frac{W}{g}\ddot{x} = W - (W + kx) - c\dot{x}. \tag{17}$$

Rearranging (17), the winding motion with damping factor is given by

$$\frac{W}{g}\ddot{x} + c\dot{x} + kx = 0, \tag{18}$$

where c denotes the damping factor, constant and determined by initial condition. Simplification of (18) yields

$$\ddot{x} + \frac{gc}{W}\dot{x} + \frac{gk}{W}x = 0, \tag{19}$$

$$\ddot{x} + 2\beta\dot{x} + \alpha^2 x = 0, \quad \frac{gc}{W} = 2\beta. \tag{20}$$

Equation (20) is a linear differential equation. One solution for (20) can be given by

$$x = e^{\mu t}, \tag{21}$$

where t is the time, and μ is obtained through (22).

$$\mu^2 + 2\beta\mu + \alpha^2 = 0, \tag{22}$$

from which

$$\mu = -\beta \pm \sqrt{\beta^2 - \alpha^2}. \tag{23}$$

For $\alpha^2 > \beta^2$, the square root value in (23) will become negative; therefore, α^2 - β^2 shows a positive value in (24).

$$\alpha_1^2 = \alpha^2 - \beta^2 \tag{24}$$

In this form, the roots of equation (22) are obtained as

$$\mu_1 = -\beta + j\alpha_1, \quad \mu_2 = -\beta - j\alpha_1. \tag{25}$$

Hence, two particular solutions for (21) can be calculated as

$$\begin{aligned} x_1 &= \frac{C_1}{2}\left(e^{\mu_1 t} + e^{\mu_2 t}\right) = C_1 e^{-\beta t}\cos\alpha_1 t, \\ x_2 &= \frac{C_2}{2j}\left(e^{\mu_1 t} - e^{\mu_2 t}\right) = C_2 e^{-\beta t}\sin\alpha_1 t, \end{aligned} \tag{26}$$

Finally, the winding displacement in the second scenario is given as

$$x = e^{-\beta t}\left(C_1\cos\alpha_1 t + C_2\sin\alpha_1 t\right). \tag{27}$$

For $\alpha^2 < \beta^2$, the roots in equation (23) will become real and negative and displacement is obtained as,

$$x = C_1 e^{\mu_1 t} + C_2 e^{\mu_2 t}. \tag{28}$$

Equation (28) shows that for $\alpha^2 < \beta^2$, there is no periodical term in winding movement. Thus, the winding will not oscillate.

C. FORCED VIBRATION WITHOUT DAMPING FACTOR

In practice, the excitation or load currents are flowing through the transformer winding and creating magnetic flux permanently. Therefore, electromagnetic force initiated by the current and magnetic flux is imposing over the transformer winding while the transformer is in service. Thus, in the absence of damping factor, the winding vibration equation is obtained as

$$\frac{W}{g}\ddot{x} = W - (W + kx) + C\sin\omega' t, \tag{29}$$

where C represents the force magnitude and ω' is the force function frequency. Differential equation of the winding displacement is then given by

$$\ddot{x} + \alpha^2 x = q\sin\omega' t, \quad q = \frac{Cg}{W}. \tag{30}$$

The solution of (30) is obtained as,

$$x_f = \eta\sin\omega' t, \quad \eta = \frac{q}{\alpha^2 - \omega^2}, \tag{31}$$

where x_f is the forced displacement. Therefore taking (15) into account, total displacement under force vibration for transformer winding without damping factor can be obtained as

$$x = x_1 + x_2 = C_1\cos\alpha t + C_2\sin\alpha t + \frac{q}{\alpha^2 - \omega^2}\sin\omega' t, \tag{32}$$

where, α comes from natural frequency of damping and ω' comes through an external force such as accurately demonstrated as

$$x = x_1 + x_2 = C_1\cos 2\pi f_n t + C_2\sin 2\pi f_n t + \frac{q}{\alpha^2 - \omega^2}\sin\omega' t, \tag{33}$$

where f_n is the natural frequency of vibration.

D. FORCED VIBRATION WITH DAMPING FACTOR

In a real case, an external force is permanently available over the transformer winding and disk spacers and barriers play as damping components reacting opposite to the winding movement. In this form, the transformer winding movement is obtained as

$$\frac{W}{g}\ddot{x} = W - (W + kx) + C\sin\omega' t - c\dot{x}. \tag{34}$$

Arrangement of (34) in standard form of differential equation yields

$$\frac{W}{g}\ddot{x} + 2\beta\dot{x} + \alpha^2 x = q\sin\omega' t. \tag{35}$$

A solution for (35) is offered as

$$x_1 = U\sin\omega' t + V\cos\omega' t. \tag{36}$$

U and V are two independent equations and are given by

$$\begin{aligned} V\omega'^2 - 2U\omega'\beta - V\alpha^2 &= 0, \\ U\omega'^2 + 2V\omega'\beta - U\alpha^2 &= \tfrac{-Cg}{W}. \end{aligned} \tag{37}$$

The solution of (37) is

$$U = \frac{Cg}{W}\left(\frac{\alpha^2-\omega'^2}{(\alpha^2-\omega'^2)^2+4\beta^2\omega'^2}\right),$$
$$V = \frac{-Cg}{W}\left(\frac{2\beta\omega'}{(\alpha^2-\omega'^2)^2+4\beta^2\omega'^2}\right). \quad (38)$$

Therefore, transformer winding displacement due an external force is obtained by

$$x = e^{-\mu t}\left(C_1\cos\alpha_1 t + C_2\sin\alpha_1 t\right) + U\sin\omega' t + V\cos\omega' t, \quad (39)$$

Equation (39) has two main terms, damping term and oscillatory term. The first term which is damping term will be diminished very fast; however, second term will cause winding oscillation permanently. The winding movement velocity is then obtained as

$$\begin{aligned}\dot{x} = &-\mu e^{-\mu t}\left(C_1\cos\alpha_1 t + C_2\sin\alpha_1 t\right)\\ &+ e^{-\mu t}\left(-C_1\alpha_1\sin\alpha_1 t + C_2\alpha_1\cos\alpha_1 t\right)\\ &+ U\omega'\cos\omega' t - V\omega'\sin\omega' t\end{aligned} \quad (40)$$

and, ultimately acceleration equation is

$$\begin{aligned}\ddot{x} = &\mu^2 e^{-\mu t}\left(C_1\cos\alpha_1 t + C_2\sin\alpha_1 t\right)\\ &- \mu e^{-\mu t}\left(-C_1\alpha_1\sin\alpha_1 t + C_2\alpha_1\cos\alpha_1 t\right)\\ &- \mu e^{-\mu t}\left(-C_1\alpha_1\sin\alpha_1 t + C_2\alpha_1\cos\alpha_1 t\right)\\ &+ e^{-\mu t}\left(-C_1\alpha_1^2\cos\alpha_1 t - C_2\alpha_1^2\cos\alpha_2 t\right)\\ &- \omega'^2\left(U\sin\omega' t + V\cos\omega' t\right).\end{aligned} \quad (41)$$

Simplification of (40) yields

$$\begin{aligned}\ddot{x} = e^{-\mu t}&\begin{pmatrix}\left(-C_1\alpha_1^2 - 2C_2\mu\alpha_1 + \mu^2 C_1\right)\cos\alpha_1 t\\ +\left(-C_2\alpha_1^2 - 2C_1\mu\alpha_1 + \mu^2 C_2\right)\sin\alpha_1 t\end{pmatrix}\\ &- \omega'^2\left(U\sin\omega' t + V\cos\omega' t\right).\end{aligned} \quad (42)$$

In final form, transformer winding vibration is

$$\begin{aligned}\ddot{x} = e^{-\mu t}&\begin{pmatrix}\left(-C_1\alpha_1^2 - 2C_2\mu\alpha_1 + \mu^2 C_1\right)\cos\alpha_1 t\\ +\left(-C_2\alpha_1^2 - 2C_1\mu\alpha_1 + \mu^2 C_2\right)\sin\alpha_1 t\end{pmatrix}\\ &- \omega'^2 P\sin(\omega' t + \varphi),\end{aligned} \quad (43)$$

where P is the magnitude coefficient of forced vibration and is obtained via $U = P\times\cos(\varphi)$, or $V = P\times\sin(\varphi)$.

From vibration modeling of the transformer, it becomes clear that the transformer core and winding vibration magnitudes are proportional to the transformer excitation voltage square and the transformer current square values, respectively. In addition, core and winding are both vibrating with fundamental frequency of 2ω, where ωmentioned to be the voltage and current signal fundamental frequencies. Therefore, theoretically, it is expected that we obtain approximately a sinusoidal time-series with frequency of 2ω for core and winding vibrations once the transformer is in service. However, this is not true in practice and higher harmonic orders ($> 2\omega$) are appeared in transformer vibration spectrum. This is most commonly experienced in distribution transformers.

For core vibration, we discussed the Lorentz force can initiate higher harmonic orders in vibration spectra, and these forces are randomly changed once transformer core materials and grain-orientation are changed.

For transformer winding, higher harmonic orders in vibration spectrum are not created by winding characteristic naturally, though they can be generated due to physical interaction between transformer core and winding; that is to say, the core vibration is transferred mechanically to the winding causing that higher harmonic orders are created by oscillation modulation. Furthermore, in oil-filled transformer, there is no access to the transformer active part directly; hence, vibration signal is taken by mounting vibration sensor over the transformer tank. This means a mixture of core and winding vibrations are recorded. Therefore, a precise interpretation of transformer vibration signal through analytical approach becomes unrealistic as higher harmonics are mixed with fundamental harmonics in vibration spectra. Furthermore, each and every transformer has its own vibration pattern, thus it is virtually impossible to achieve a comprehensive analytical model for vibration analysis. Moreover, an analytical model cannot be utilized for prognosis. A plausible approach to address intractability of analytical models is to develop predictive data-driven models of transformer fault detection. To lay our groundwork for developing effective prognostic models of transformer conditions based on vibration signals, we consider and conduct two practical studies as described in the next section.

FIGURE 3. 15 kVA cast-resin three-phase 400/400 V transformer

IV. CASE STUDY

A. CASE STUDY 1: (TRANSFORMER UNDER AND OVER EXCITATIONS)

1) MEASUREMENT

The first practical study was conducted on a three-phase distribution 400/400 V, 15 kVA cast-resin transformer to emulate over-excitation study, see Fig. 3. In this study, the primary side of transformer was connected to a three-phase power supply and secondary side was connected to 15 kW resistive load, PF = 1. Excitation voltage in the primary side of transformer was changed from 80 % to 110 % with steps of 5% using a three-phase autotransformer, see Table 1. Vibration sensors were mounted on transformer core and winding top to collect vibration signal for both transformer core and winding, see Fig. 3. Vibration signal in time domain for each

TABLE 1. No-load transformer excitation voltages.

Excitation Voltage %	80%	85 %	90 %	95 %	100 %	105 %	110 %
Injected Voltage [V]	320	340	360	380	400	420	440

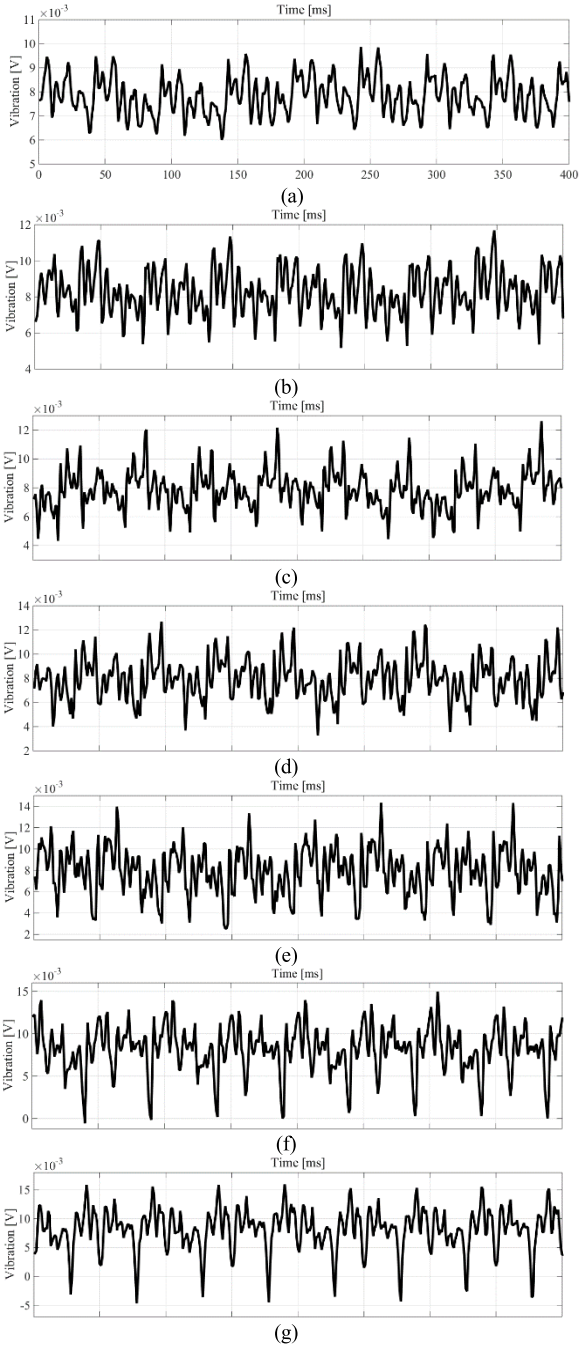

FIGURE 4. Time-series core vibration signals, (a) 80% excitation, (b) 85% excitation, (c) 90% excitation, (d) 95% excitation, (e) 100% excitation, (f) 105% excitation, (g) 110% excitation.

step was recorded and illustrated in Fig. 4. Predictive model development is discussed in the next subsection.

2) SIGNAL PROCESSING

In this case study, we aimed to construct a predictive model of excitation voltage based on vibration waveforms. The continuous nature of excitation (or injected) voltage suggests constructing a regression model. Our predictive analysis is based on four steps: (1) *Signal Modeling*: we first model the vibration time-series using a sum of sinusoids of order M. Using this modeling, we are able to represent the vibration time-series in a much lower dimensional subspace spanned by a set of attributes; (2) *Regression Model Construction and Model Order Selection*: the attributes found in step (1) are used to construct the regression model. The order of the signal model M used in step (1) is the one that results in the lowest error rate of the constructed regression model determined by a cross-validation procedure; and (3) *Validation*: the error rate of the final constructed regression model is determined on the test data. Therefore, at end of this process, we plan to predict transformer under- or over-excitation in any condition using vibration signal. It should be noted that according to the analytical model, transformer voltage excitation will influence core vibration; thus, we focused on core vibration signal in this part.

3) SIGNAL MODELING

Based on the periodic nature of core vibration measurements, we choose the sum of sinusoids signal model with unknown amplitudes, frequencies, and phases [22]. We use this mathematical form to model the core vibration time-series on short observation intervals of 0.1 sec during which we model the signal as described in this section. This divides the T sec long vibration time-series to $10T$ individual observation vectors that can be used for training the regression model (see Table 2).

TABLE 2. Number of training observation vectors for each injected voltage level.

Number of training observation vectors	481	329	286	394	432	476	462
Injected Voltage [V]	320	340	360	380	400	420	440

Denoting the signal by $s[n]$, we have

$$\begin{aligned} s[n] &= \sum_{i=1}^{M} A_i cos(2\pi f_i n + \varphi_i) \\ &= \sum_{i=1}^{M} \alpha_{1i} cos(2\pi f_i n) + \alpha_{2i} sin(2\pi f_i n), \\ n &= 0, 1, \ldots, N-1, \end{aligned} \tag{44}$$

where M is the model order, $\alpha_{1i} = A_i \cos(\varphi_i)$, $\alpha_{2i} = -A_i \sin(\varphi_i)$, $0 < f_i < f_{i+1} < 0.5$, $\forall i$, and the second equality in (44) is the result of regular trigonometric identities. The sinusoidal signal model presented in (44) is linear in terms of unknown parameters α_{1i} and α_{2i}, and nonlinear in terms of f_i, $i = 1, \ldots, M$. Without any distributional assumption on the observations or additive noise, we follow a nonlinear least squares (NLS) estimation procedure [23] to estimate the parameters as described next. We first estimate f_i's by a grid search over $\{0.005, 0.01, \ldots, 0.495\}^M$. In particular,

the NLS estimates of frequencies are obtained by maximizing the following function [23]:

$$\left[\hat{f}_1, \ldots \hat{f}_M\right] = \underset{f_1,\ldots f_M}{\operatorname{argmax}} J\left(f_1, \ldots, f_M\right) \triangleq \boldsymbol{x}^{\mathbf{T}} \boldsymbol{H}\left(f_1, \ldots, f_M\right)\left(\boldsymbol{H}^T\left(f_1, \ldots, f_M\right) \times \boldsymbol{H}\left(f_1, \ldots, f_M\right)\right)^{-1} \boldsymbol{H}^T\left(f_1, \ldots, f_M\right) \boldsymbol{x} \quad (45)$$

where $\boldsymbol{x}$ is a $N \times 1$ observation vector (a sampling rate of 3 kHz leads to $N = 300$), and $H\left(f_1, \ldots, f_M\right)$ is an $N \times 2M$ observation matrix with the element on the i^{th} row and j^{th} column determined as

$$\boldsymbol{H}\left(f_1, \ldots, f_M\right)[i, j] = \begin{cases} \cos\left(2\pi f_{\lceil j/2 \rceil}\left(i-1\right)\right), & for\ j\ odd \\ \sin\left(2\pi f_{\lceil j/2 \rceil}\left(i-1\right)\right), & forj\ even, \end{cases} \quad (46)$$

where $\lceil a \rceil$ denotes the least integer greater or equal to a. After estimating f_i's, the linear parameters are estimated as

$$\hat{\boldsymbol{\alpha}} \triangleq \left[\hat{\alpha}_{11}\ \hat{\alpha}_{21} \ldots \hat{\alpha}_{M2}\right]^T = \left(\boldsymbol{H}^T\left(\hat{f}_1, \ldots, \hat{f}_M\right) \boldsymbol{H}\left(\hat{f}_1, \ldots, \hat{f}_M\right)\right)^{-1} \boldsymbol{H}^T\left(\hat{f}_1, \ldots \hat{M}\right) \boldsymbol{x}, \quad (47)$$

where $\hat{\boldsymbol{\alpha}}$ denotes the estimator of parameter $\boldsymbol{\alpha}$.

4) REGRESSION MODEL CONSTRUCTION AND MODEL-FEATURE SELECTION

In this work, we construct four regression models in order to predict real-valued excitation voltage: 1) linear regression; 2) model trees; 3) support vector regression with Gaussian kernel, also known as radial basis function (SVR-RBF) [28]; and 4) multilayer perceptron (MLP) [28].

Linear regression lends itself to the ease of interpretability and accuracy once the response variable can be explained by a linear combination of predictors. On the other hand, a model tree is a piece-wise linear regression model and has been originally proposed and implemented by M5 algorithm [24] and its modification known as M5' [25]. While regression trees such as CART have real-valued at their leaves, model-trees fit a multivariate linear regression function to the data at each leaf node to predict the continuous response variable. It has been shown that not only can model trees be learned very efficiently, but they can also outperform regression trees [24], [26].

The attributes used in our predictive regression models are primarily determined by modeling vibration signals as described in the previous section (Signal Modeling). In particular, in (44) we consider models of order $M = 2, 3, 4$ which lead to 6, 9, and 12 attributes, respectively ($2M$ parameters in $\hat{\boldsymbol{\alpha}}$ defined in (47) and M frequency parameters). Nevertheless, for a model of order M, not all attributes are necessarily important. That is to say, the model selection stage contains a natural feature selection procedure to determine the subset of $3M$ features with the best predictive capacity. Therefore, in the intertwined model-feature selection stage, we need to determine: (1) the order M that is used to model vibration signals (model selection); (2) the subset of $3M$ features (feature selection); and (3) the structure of the model (linear regression, model tree, SVR-RBF, or MLP) that uses the feature subset and leads to the lowest error rate (model selection). The latter stage also includes choosing the variance of kernel σ^2 in SVR-RBF, and the size and number of hidden layers in MLP.

We use a nested 5-fold cross-validation procedure to select the model (both M and the structure) and feature subset that lead to the lowest error rate of the predictive model on training data. In K-fold cross-validation (CV), one randomly divides the set of training data to K subsets, successively holds out these subsets from the training data, construct the predictive model on the reduced training dataset, and determine the error rate of the constructed (surrogate) predictive models on the held-out subset [27]. The error rate of the constructed predictive model on the full training data is then the average of the error rates of the surrogate regression models. Nevertheless, to avoid the bias selection, which results in an overly optimistic error rate of the final constructed predictive model [28], it is essential to apply the cross-validation procedure *external* to feature selection. At the same time, in the feature selection stage, we use an *internal* cross-validation to evaluate the predictive capacity of the predictive models constructed using each subset of features (this known as wrapper feature selection [29]). As the search strategy to find the best possible feature subset, we conducted an exhaustive search of the feature space. That is to say, we construct 2^{3M}-1 model for $M = 2, 3, 4$, and choose the subset that possesses the lowest error rate. Figure 5 provides a schematic description of this nested cross-validation procedure. As the measure of error rate to rank all the feature subsets, we use the Mean Absolute Error (MAE):

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |pv_i - av_i|}{n} \quad (48)$$

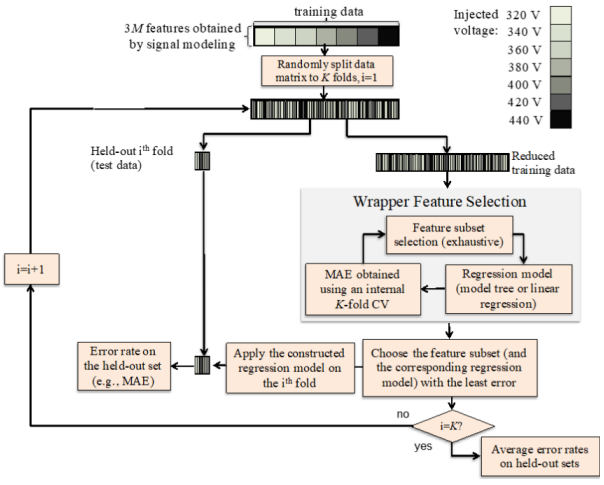

FIGURE 5. A schematic diagram of nested K-fold cross-validation with wrapper feature selection.

where n is the number of test data points (e.g., the size of each held-out set in cross-validation), pv_i and av_i are the predicted and actual values on the test instances, respectively. For the SVR-RBF, we assumed $\sigma^2 \in \{0.01, 0.02, 0.1, 1, 10, 50, 100\}$ and chose the one that led to the lowest MAE in the proposed cross-validation procedure.

A similar strategy was implemented to choose the number of hidden layers and neurons in MLP. To guard against overfitting, we only considered a MLP with a single or two hidden layers [30]. Denoting the number of initial features by p, we varied the total number of neurons used in the model in $\{p/2, p, 2p, 3p\}$ and chose the structure with the lowest MAE. When two hidden layers were considered, we used the same number of neurons in both layers. This set of possible number of neurons has been chosen to cover typical suggestions in the literature about the size of hidden layers. For example, Boger and Guterman suggest taking the size of hidden layer as the number of dimension p [31]. Nevertheless, according to Kolmogorov theorem, any function of p variables can be represented by a superposition of $2p + 1$ univariate functions [32]. Referring to this universal approximation theorem, Swingler [32] suggests that in a MLP with a single hidden layer, we do not need a number of neurons larger than twice the number of dimension (i.e., $\leq$2p) [32, p. 53]. Interestingly, in most of our experiments, the performance of the constructed MLP has been optimized at $2p$ total number of neurons.

The aforementioned model selection procedure suggests that for $M = 2$, the attributes of the signal model used in the SVR-RBF leads to the lowest MAE of the constructed regression models. Using other performance measures such as, Relative Absolute Error (RAE), and Root Relative Squared Error (RRSE) defined as [33]

$$\text{RAE} \triangleq 100 \times \frac{\sum_{i=1}^{n} |pv_i - av_i|}{\sum_{i=1}^{n} |av_i - \overline{av}|}, \tag{49}$$

$$\text{RRSE} \triangleq 100 \times \sqrt{\frac{\sum_{i=1}^{n} |pv_i - av_i|^2}{\sum_{i=1}^{n} |av_i - \overline{av}|^2}}, \tag{50}$$

has also led to a similar conclusion (using $M = 2$ and SVR-RBF). Table 3 shows these performance measures for all models and $M = 2, 3, 4$.

5) VALIDATION

As an additional validation step, we have examined the performance of our constructed SVR-RBF in predicting excitation voltage in a set of vibration time-series that has been collected independently from training data. Table 4 shows the number of test observation vectors (of duration 0.1 sec) for each injected voltage level. Table 5 presents the performance measures of the constructed SVR-RBF. The results of this table confirms that the constructed regression model is remarkably accurate in predicting real-valued excitation voltage. This result shows the nonlinearity of excitation voltage in terms of model parameters used in (44).

Table 5 shows that the trained model is able to recognize transformer excitation voltage using recorded signal in last 0.1 s. The error for MAE, RAE, RRSE are quite reasonable $< 10\%$. Therefore, using transformer core vibration database, it is smart to find excitation voltage growing trend and indicate undesirable transformer overvoltage in less than a second.

TABLE 3. Model selection: SVR-RBF and $M = 2$ leads to smallest error rate measured by various performance metrics in a cross-validation procedure. The estimates of the variance of SVM-RBF and the number/size of hidden layers in MLP that led to lowest MAE are identified as "[.]".

Model Tree	M=2	M=3	M=4
MAE (Model Tree)	3.32	1.92	2.11
RAE (Model Tree)	9.18%	5.30%	5.84%
RRSE (Model Tree)	12.35%	8.05%	10.22%
MAE (Linear Regression)	28.22	34.24	19.54
RAE (Linear Regression)	77.96%	94.60%	53.98%
RRSE (Linear Regression)	79.31%	98.30%	58.26%
MAE, SVR-RBF [σ^2 (variance of kernel) =]	**0.89** **[0.1]**	1.73 [1]	2.48 [1]
RAE, SVR-RBF [σ^2 (variance of kernel) =]	**2.47%** **[0.1]**	4.79% [1]	6.77% [1]
RRSE, SVR-RBF [σ^2 (variance of kernel) =]	**3.64%** **[0.1]**	9.55% [1]	10.38% [1]
MAE, MLP [number/size of hidden layer(s) =]	8.87 [1/12]	2.68 [1/18]	1.66 [2/12]
RAE, MLP [number/size of hidden layer(s) =]	24.49% [1/12]	7.40% [1/18]	4.60% [2/12]
RRSE, MLP [number/size of hidden layer(s) =]	27.79% [1/12]	10.56% [1/18]	6.72% [2/12]

TABLE 4. Number of test observation vectors for each injected voltage level.

Number of training observation vectors	160	109	95	131	143	158	153
Injected Voltage [V]	320	340	360	380	400	420	440

TABLE 5. Performance measures of the constructed SVR-RBF on independent test observations.

MAE	RAE	RRSE
0.87	%2.42	3.24%

B. CASE STUDY 2: (INTER-TURN SHORT-CIRCUIT STUDY)

1) MEASUREMENT

The second practical study was conducted on a smaller transformer to analyze vibration data and recognize automatically the transformer inter-turn fault by signal processing. In this process, when the transformer was in service, the transformer winding vibration under different loads was used to train a model. Afterwards, this developed model was utilized to predict the transformer overload and inter-turn short circuit; that is to say, this part is focused on developing an algorithm which can be programmed over the cloud system and predict transformer turn-to-turn insulation failure in early stages (prognosis). Examination and analysis were performed on an open wounded 240/30/30 V, 252 VA single-phase transformer.

TABLE 6. Different loads for transformer which are used for training the predictive model using winding vibration signals.

Load [p.u]	0.71	0.95	1	1.19	1.31	1.42	1.66
Load current [A]	3	4	4.2	5	5.5	6	7
Load [p.u]	1.9	2	2.14	2.26	2.38	2.61	2.85
Load current [A]	8	8.4	9	9.5	10	11	12

The nominal current for the transformer primary side was 1.05A and for the secondary side 4.2A for each winding. To run the test setup, in the first stage, the primary side of this transformer (winding A) was connected to the voltage source, 240 V, and secondary side (winding B) connected to a variable resistive load. The load was changed step by step and different transformer winding vibration signals were recorded. Variable loads and steps used for training our predictive model are listed in Table 6. In fact, in practice, the load varies over the transformer daily or hourly and it causes different vibrations over transformer winding and we can use this transformer real-time winding vibration under different loads to train the model. Two vibration sensors were used to monitor transformer vibration. One sensor was mounted on the transformer core top, one sensor over the transformer winding top. Vibration signals in time domain for winding and core of transformer were recorded and transferred directly to the cloud environment for calculation and analysis. However, winding vibration was used in this part as we focused on winding current and its short-circuit.

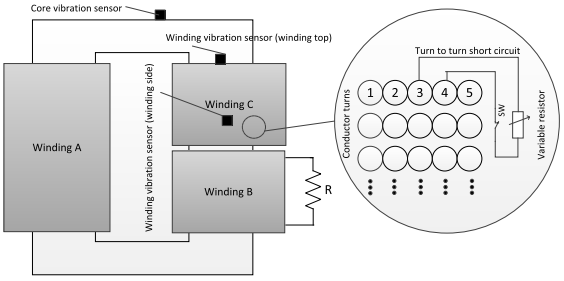

FIGURE 6. Single-phase transformer inter-turn fault emulation schematic.

TABLE 7. Transformer short-circuit current used for testing the predictive capacity of the constructed model.

Load [p.u]	3.09	3.57
Short-circuit current [A]	13	15

Afterwards, to examine and validate the trained model over the terminals of third winding (winding C), turn-to-turn short circuit was practically emulated by connecting two conductors of winding C together. We put a variable resistor between conductors to control the short circuit current, see Fig. 6. Short-circuit current was emulated for 13 and 15 A, see Table 7. Figure 7 shows the test setup, sensors, loads, and portable device and Figure 8 illustrates recorded vibration signal from winding top sensor.

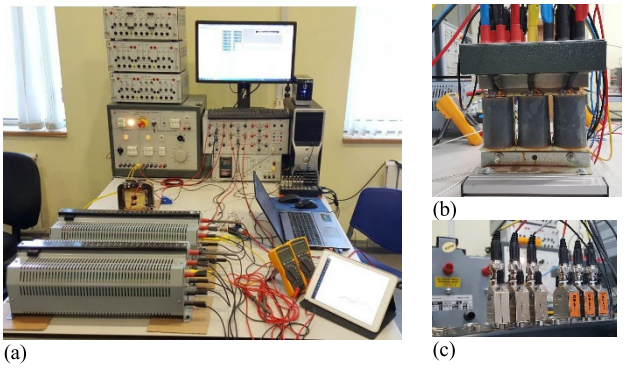

FIGURE 7. Practical test set-up for single-phase three-winding test object.

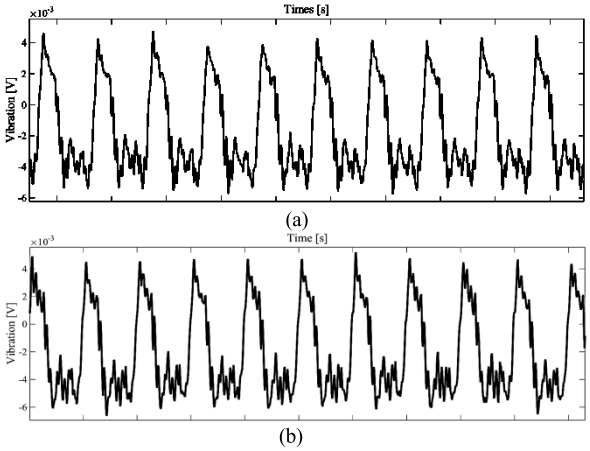

FIGURE 8. Time-series winding top vibration signal due to short-circuit, (a) 13 A short-circuit current, (b) 15 A short-circuit current, *x axis* = 400 ms.

2) SIGNAL PROCESSING

We use a similar signal model and estimation strategy as described in (44)-(47). Similar to Section 2.2, we developed four types of models (linear regression, model tree, SVR-RBF, MLP) to predict the transformer overloads. However, we choose a strategy different from cross-validation in the model selection stage. In order to select the model (the order of the model M and the structure of models), we first developed our models on load currents ranging from 3 to 10A, and examined their performance on observation for the load currents of 11 and 12A. This validation strategy is used to simulate real scenarios where training the predictive model is conducted on a set of observations collected from a non-overloaded transformer to predict the transformer overload and inter-turn short circuit (see Fig. 9).

In our case, we collect training observations for the load currents ranging from 3 to 12A and examine the performance of the constructed model in predicting overload currents 13 and 15A, which we assume do not occur during the training stage. Table 8 provides various performance measures obtained using this model selection strategy for all models. This procedure suggests that for $M = 2$, the attributes of

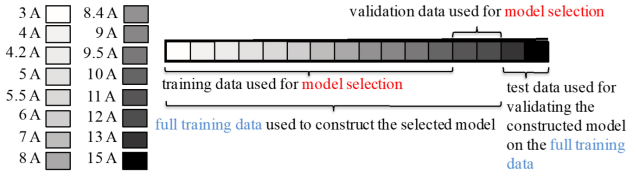

FIGURE 9. The model selection phase is conducted using observations for load currents ranging from 3 to 10A and validated on observations of 11 and 12 A. Once the model is selected, the full training data (observations ranging from 3 to 12A) are used to construct the model. This constructed model is validated on observations ranging for load currents of 13 and 15A.

TABLE 8. Model selection: MLP with 1 hidden layer and 8 neurons for $M = 2$ leads to smallest error rate measured by various performance metrics. We validate both the constructed MLP and the linear regression model. The linear regression model is chosen for the ease of its interpretability while having a comparable performance with other models. The estimates of the variance of SVM-RBF and the number/size of hidden layers in MLP that led to lowest MAE are identified as "[.]".

Model Tree	M=2	M=3	M=4
MAE, Model Tree	1.92	1.91	1.84
RAE, Model Tree	33.32%	33.22%	32.13%
RRSE, Model Tree	33.91%	33.84%	32.86%
MAE, Linear Regression	1.33	1.14	**1.12**
RAE, Linear Regression	23.20%	19.82%	**19.50%**
RRSE, Linear Regression	23.41%	20.25%	**19.84%**
MAE, SVR-RBF [σ^2 (variance of kernel) =]	1.03% [10]	1.01% [1]	1.30% [10]
RAE, SVR-RBF [σ^2 (variance of kernel) =]	16.81 % [10]	16.62% [1]	22.12% [10]
RRSE, SVR-RBF [σ^2 (variance of kernel) =]	17.35% [10]	17.27% [1]	22.57% [10]
MAE, MLP [number/size of hidden layer(s) =]	**0.98%** **[1/8]**	1.87% [1/12]	1.78% [2/2]
RAE, MLP [number/size of hidden layer(s) =]	**15.73%** **[1/8]**	31.66% [1/12]	29.94% [2/2]
RRSE, MLP [number/size of hidden layer(s) =]	**16.98%** **[1/8]**	32.41% [1/12]	30.85% [2/2]

the signal model used in the MLP structure leads to the lowest MAE (and the lowest RAE and RRSE). After model selection, we use the full training data for the load currents ranging from 3 to 12 A to train our regression models.

The constructed MLP and the linear regression model are both validated on the observations for a set of short-circuit currents (13 and 15A) from which we had no observation in the training stage (short-circuit current can take any current flowing through the circuit but we selected and set the short-circuit setup on 13 and 15A as it was more realistic and precise to examine the trained model with unseen currents in the circuit). The MLP is chosen for validation as it possesses the lowest error rate in the training stage and the linear regression model is taken because it has a remarkably simple structure while having a comparable performance with other models.

Table 9 provides various performance measures of these constructed models on this set of test data. Interestingly, the constructed linear model presented in (51) shows even a better performance than the constructed MLP in predicting the overload currents. Nevertheless, these low error

TABLE 9. Performance measures of the constructed MLP and linear regression model (51) on a set of independent observations from overload currents 13 and 15 A that have not been measured during the model training stage. Interestingly, linear regression shows a better performance on the test observations.

Model	MAE	RAE	RRSE
Linear Regression (M=4)	0.78	11.37%	12.19%
MLP (M=2, size of hidden layer 8)	1.25	18.31%	21.71%

rates (20%) suggest the efficacy of both constructed models in predicting overload currents. In other words, the constructed models are capable of precisely recognizing transformer short-circuit current (or any undesirable load current) by means of recorded vibration signal in the last 0.1 s observation window. Therefore, the use of transformer winding vibration data has enabled us to construct a smart prognosis system to detect inter-turn short-circuit fault in early stage.

It would be interesting to look into the structure of the simple constructed linear regression model. There are four variables in this model that are linearly combined to predict the response variable:

$$y = c + \beta_1 A_1 + \beta_2 A_2 + \beta_3 A_3 + \beta_4 f_4, \qquad (51)$$

where y is the load current, A_i, $i = 1, 2, 3$, denotes the amplitude of the sinusoids with frequency f_i used in the model defined in (44), f_4 is the largest frequency used in this signal model (recall that in (44) we have $f_i < f_{i+1}$), and c and β_i are presented in Table 10. This model suggests the amplitude of the first three sinusoids along with the frequency of the fourth sinusoids are the only factors needed to predict the overload current.

TABLE 10. Intercept and coefficients of the predictors used in the linear regression model (51).

c	β_1	β_2	β_3	β_4
-8.18	3322.70	4776.45	2790.3	-18.63

V. CLOUD COMPUTING

Internet of Things (IoT) has advantages over local data monitoring center in transformer prognosis in terms of factors such as dynamic monitoring, reliability and availability, worldwide 24-hour care and visibility, reduction of monitoring system costs and simple end user access to data. Having said that, vibration data analysis for transformer fault detection before failure, which is known as prognosis is performed using cloud computing. Using abovementioned analytical calculation and signal processing techniques, appropriate libraries and algorithms is setup over the cloud environment and a protocolled portable device (any tablet or smartphone) is connected to the cloud environment to monitor and access data. Developed algorithm over the cloud system is able to monitor entire vibration signals, perform real-time analysis and act properly, e.g., notify operator as to undesirable condition in transformer. This system is also capable of activating a protection relay and disconnect transformer from main

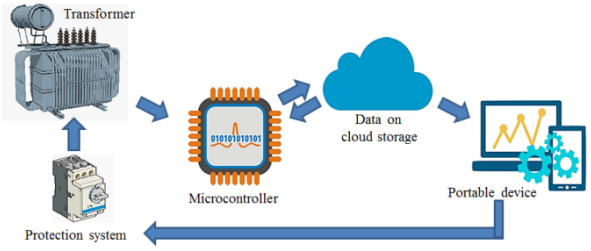

FIGURE 10. Operational principle of the system, a general view.

feeder once short-circuit fault current is initiated and the vibration spectrum becomes undesirable. Figure 11 illustrates an overview of transformer vibration analysis and operation principle using cloud computing in this study.

Real-time monitoring of transformer using IoT requires three connections to be ensured. Microcontroller should be connected to the cloud platform; cloud platform to portable device application; and, portable device to the protection system of transformer. In this study, programmed microcontroller is connected to the IoT platform through the local network Ethernet connection. The initial task for this step is to design a specific hardware setup model. The next step is to provide a gateway between hardware and cloud system which enables data to reach cloud service. One of the feasible and reliable gateways is MQ Telemetry Transport (MQTT) broker-based publish subscribe messaging protocol which runs as a broker between device and the cloud system. MQTT broker is to help and identify proper connection between microcontroller device and cloud platform. MQTT protocol device credential is a function of how the microcontroller platform sends data to cloud system.

Synchronizing IoT platform with the Ethernet shield of microcontroller platform is crucially vital. Therefore, microcontroller platform is established according to libraries, device credentials, and Ethernet shield mac address. Device password is also registered in cloud platform. Two main advantages of the IoT platform utilized in this study are: (1) it is asynchronous with various levels of quality service, which is important where internet connections become unreliable; and (2) it does not require a memory-consuming software for a client to get connected, which makes it suitable for devices with limited memory. With MQTT, information is assigned with a specific message unit and sign in both sending and receiving points. This procedure helps identify accurate information. Then the system starts to work on the connection between application and cloud. The Bluemix account is used and a new mobile application project is designed in dashboard console (including the services as notification) to keep user engaged and Cloudant NoSQL DB is used to provide the access to JSON data layer.

In addition, Node-RED which is a programming tool organizes the connection between cloud and application. It helps establish multi-connection clouding system and connecting other devices. This is important as it is planned to send data from cloud to android application. APIs of the cloud and online services are also used for establishing connection and data gathering.

VI. CONCLUSION

Transformer core and winding vibrations were mathematically modeled and their acceleration factors were analytically discussed in detail. It was technically discussed that using vibration time-series, the transformer vibration analytical model is unskilled to interpret abnormal and faulty transformer conditions and contribute to transformer prognosis. In addition, in the context of Industry 4.0, a prognosis technique becomes meaningful if it can be implemented and developed over a cloud system and work based on IoT protocols.

In this work, we utilized the state-of-the-art machine learning and signal processing techniques to analyze transformer vibration signals and train a real-time model that can predict transformer turn-to-turn short circuit fault based on the most recent 0.1 s duration vibration measurements. This technique was also able to detect transformer over and under excitations. Two test objects were practically examined through introduced vibration analysis technique. The data obtained from practical studies were transferred to the cloud and the developed vibration signal processing technique for real-time prognosis became accessible over any portable device. For future study, it is planned to connect the cloud system to a protection device and control or trip transformer before catastrophic failure.

REFERENCES

[1] M. Bagheri, M. S. Naderi, and T. Blackburn, "Advanced transformer winding deformation diagnosis: Moving from off-line to on-line," *IEEE Trans. Dielectrics Electr. Insul.*, vol. 19, no. 6, pp. 1860–1870, Dec. 2012.

[2] M. H. Samimi and S. Tenbohlen, "FRA interpretation using numerical indices: State-of-the-art," *Int. J. Electr. Power Energy Syst.*, vol. 89, pp. 115–125, Jul. 2017.

[3] M. F. M. Yousof, C. Ekanayake, and T. K. Saha, "Frequency response analysis to investigate deformation of transformer winding," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 22, no. 4, pp. 2359–2367, Aug. 2015.

[4] M. Bagheri, S. Nezhivenko, and B. T. Phung, "Loss of low-frequency data in on-line frequency response analysis of transformers," *IEEE Elect. Insul. Mag.*, vol. 33, no. 5, pp. 32–39, Sep./Oct. 2017.

[5] Z. Zhao, C. Yao, C. Li, and S. Islam, "Detection of power transformer winding deformation using improved FRA based on binary morphology and extreme point variation," *IEEE Trans. Ind. Electron.*, vol. 65, no. 4, pp. 3509–3519, Apr. 2018.

[6] R. Rajamani, M. Rajappa, K. Arunachalam, and B. Madanmohan, "Inter-turn short diagnosis in small transformers through impulse injection: On-line on-load self-impedance transfer function approach," *IET Sci. Meas. Technol.*, vol. 11, no. 8, pp. 961–966, 2017.

[7] M. Bagheri, M. S. Naderi, T. Blackburn, and T. Phung, "Frequency response analysis and short-circuit impedance measurement in detection of winding deformation within power transformers," *IEEE Elect. Insul. Mag.*, vol. 29, no. 3, pp. 33–40, May 2013.

[8] E. Arri, A. Carta, F. Mocci, and M. Tosi, "Diagnosis of the state of power transformer windings by on-line measurement of stray reactance," *IEEE Trans. Instrum. Meas.*, vol. 42, no. 2, pp. 372–378, Apr. 1993.

[9] N. Shu, C. Zhou, F. Hu, Q. Liu, and L. Zheng, "Study on ultrasonic measurement device for transformer winding deformation," in *Proc. Int. Conf. Power Syst. Tech. (PowerCon)*, vol. 3. 2002, pp. 1401–1404.

[10] M. A. Hejazi, G. B. Gharehpetian, G. Moradi, H. A. Alehosseini, and M. Mohammadi, "Online monitoring of transformer winding axial displacement and its extent using scattering parameters and k-nearest neighbour method," *IET Generat., Transmiss. Distrib.*, vol. 5, no. 8, pp. 824–832, Aug. 2011.
[11] A. Abu-Siada and S. Islam, "A novel online technique to detect power transformer winding faults," *IEEE Trans. Power Del.*, vol. 27, no. 2, pp. 849–857, Apr. 2012.
[12] T.-T. He, J-D. Wang, J. Guo, H. Huang, X.-X. Chen, and J. Pan, "A vibration based condition monitoring system for power transformers," in *Proc. Asia–Pacific Power Energy Eng. Conf. (APPEEC)*, 2009, pp. 1–4.
[13] Y. X. Liao, T. Y. Zhu, Y. Q. Sun, J. Zhang, T. Cheng, and Y. Wang, "Load influence on lissajous figure for online transformer winding diagnosis," in *Proc. IEEE Int. Conf. High Voltage Eng. Appl. (ICHVE)*, Sep. 2016, pp. 1–4.
[14] P. M. Joshi and S. V. Kulkarni, "Transformer winding diagnostics using deformation coefficient," in *Proc. IEEE Power Energy Soc., General Meeting, Conversion Del. Electr. Energy*, Jul. 2008, pp. 1–4.
[15] B. Garcia, J. C. Burgos, and A. M. Alonso, "Transformer tank vibration modeling as a method of detecting winding deformations—Part I: Theoretical foundation," *IEEE Trans. Power Del.*, vol. 21, no. 1, pp. 157–163, Jan. 2006.
[16] J. Shengchang, L. Yongfen, and L. Yanming, "Research on extraction technique of transformer core fundamental frequency vibration based on OLCM," *IEEE Trans. Power Del.*, vol. 21, no. 4, pp. 1981–1988, Oct. 2006.
[17] M. Bagheri and B. T. Phung, "Frequency response and vibration analysis in transformer winding turn-to-turn fault recognition," in *Proc. IEEE Conf. Smart Green Technol. Electr. Inf. Syst. (ICSGTEIS)*, Oct. 2016, pp. 10–15.
[18] S. Saponara, L. Fanucci, F. Bernardo, and A. Falciani, "Predictive diagnosis of high-power transformer faults by networking vibration measuring nodes with integrated signal processing," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 8, pp. 1749–1759, Aug. 2016.
[19] S. Saponara, "Distributed measuring system for predictive diagnosis of uninterruptible power supplies in safety-critical applications," *Energies*, vol. 9, no. 5, p. 327, 2016, doi: 10.3390/en9050327.
[20] H. Zhou, K. Hong, H. Huang, and J. Zhou, "Transformer winding fault detection by vibration analysis methods," *Appl. Acoust.*, vol. 114, pp. 136–146, Dec. 2016.
[21] K. Hong, H. Huang, and J. Zhou, "Winding condition assessment of power transformers based on vibration correlation," *IEEE Trans. Power Del.*, vol. 30, no. 4, pp. 1735–1742, Aug. 2015.
[22] R. H. Shumway and D. S. Stoffer, *Time-Series Analysis and its Applications: With R Examples*, 4th ed. New York, NY, USA: Springer, 2017.
[23] S. M. Kay, *Fundamentals of Statistical Signal Processing: Practical Algorithm Development*, vol. 3. San Francisco, CA, USA: Prentice-Hall, 2013.
[24] R. J. Quinlan, "Learning with continuous classes," in *Proc. 5th Australian Joint Conf. Artif. Intell.*, Singapore, 1992, pp. 343–348.
[25] Y. Wang and I. H. Witten, "Induction of model trees for predicting continuous classes," in *Proc. 9th Eur. Conf. Mach. Learn.*, 1997, pp. 1–10.
[26] E. Frank, Y. Wang, S. Inglis, G. Holmes, and I. H. Witten, "Using model trees for classification," *Mach. Learn.*, vol. 32, no. 1, pp. 63–76, 1998.
[27] L. Devroye, L. Gyorfi, and G. Lugosi, *A Probabilistic Theory of Pattern Recognition*. New York, NY, USA: Springer, 1997.
[28] C. Ambroise and G. J. McLachlan, "Selection bias in gene extraction on the basis of microarray gene-expression data," *Proc. Nat. Acad. Sci.*, vol. 99, no. 10, pp. 6562–6566, 2002.
[29] R. Kohavi and G. John, "Wrappers for feature selection," *Artif. Intell.*, vol. 97, nos. 1–2, pp. 273–324, Dec. 1997.
[30] M. Berry and G. S. Linoff, *Data Mining Techniques*, 2nd ed. Hoboken, NJ, USA: Wiley, 2004, p. 231.
[31] Z. Boger and H. Guterman, "Knowledge extraction from artificial neural network models," in *Proc. IEEE Syst., Man, Cybern. Conf.*, Orlando, FL, USA, Oct. 1997, pp. 3030–3035.
[32] K. Swingler, *Applying Neural Networks: A Practical Guide*. San Francisco, CA, USA: Morgan Kaufmann, 1996.
[33] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed. Amsterdam, The Netherlands: Elsevier, 2016.
[34] Intel IT Center. *Solution Brief Big Data in the Cloud: Converging Technologies*, Accessed: Apr. 2015. [Online]. Available: https://www.intel.com/content/dam/www/public/us/en/documents/product-briefs/big-data-cloud-technologies-brief.pdf

MEHDI BAGHERI (M'12) received the M.Sc. degree in power engineering from the Sharif University of Technology, Tehran, Iran, in 2007, and the Ph.D. degree from the University of New South Wales, Sydney, Australia, in 2014. He joined the Iran Transformer Research Institute, Tehran, as a Research Engineer, and was the Head of the Test and Diagnostic Department, from 2008 to 2010. From 2015 to 2016, he served as a Post-doctoral Research Fellow with the Electrical Engineering Department, National University of Singapore, working closely with Rolls-Royce Pte. Ltd., on condition monitoring and predictive maintenance of marine transformers and filters. He is currently an Assistant Professor with the School of Electrical and Electronic Engineering, Nazarbayev University, Astana, Kazakhstan. His research interests include field and marine applications of high-voltage engineering, condition monitoring and diagnosis of power transformers and electrical rotating machines, transients in power systems and power quality. He is a member of the IEEE Dielectrics and Electrical Insulation Society.

AMIN ZOLLANVARI (M'10) received the B.Sc. and M.Sc. degrees in electrical engineering from Shiraz University, Iran, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, USA, in 2010. He held a Post-doctoral position with the Harvard Medical School and Brigham and Women's Hospital, Boston, MA, USA, from 2010 to 2012, and then joined the Department of Statistics, Texas A&M University, as an Assistant Research Scientist, from 2012 to 2014. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Nazarbayev University, Astana, Kazakhstan. His research interest includes signal processing and machine learning.

SVYATOSLAV NEZHIVENKO received the bachelor's degree in power engineering from Karaganda State Technical University, Karaganda, Kazakhstan, in 2009, and the M.Sc. degree in electrical and computer engineering from The University of British Columbia, Canada, in 2014. He was the Head of the Communication Department, Astana Power Plant, in 2015. He is currently pursuing the Ph.D. degree in power engineering with Nazarbayev University. His research interest includes high voltage and transformer monitoring and diagnosis.