Remaining lifetime prediction of distribution transformer based on improved hidden semi-Markov model

FangZhou Hao 1, Zili Li* 2, Haiyan Wen 2, Hongshan Zhao 2

1 Guangzhou power supply bureau, 510000, Guangzhou, China; 2 School of Electrical & Electronic Engineering, North China Electric Power University, 071003, Baoding, China

Corresponding author: 296814253@qq.com

Abstract. The distribution transformer is the key equipment in the distribution network. The state assessment and residual life prediction of the distribution transformer is an important prerequisite for the state maintenance. Based on the operating data of distribution transformers, this paper proposes a life prediction method based on time-varying hidden semi-Markov model. Firstly, the traditional hidden semi-Markov model is introduced, and the shortcomings of the method are expounded. Then, the hidden semi-Markov model with time-varying transition probability matrix is proposed for the fixed state transition probability. Finally, the power distribution is used. The operating data of the transformer is verified against the improved method. The conclusions show that the time-varying hidden semi-Markov model is more accurate than the traditional model residual life estimate.

1. Introduction

The distribution network is an important link between the power system and the user, and the distribution transformer plays an important role in it. Therefore, it is very important to improve the operation and management level of the distribution transformer for the safe operation of the entire power grid. Using modern technology and methods to evaluate and predict its remaining life, it is of great significance to achieve its maximum service life under safe conditions.[1].

At present, the estimation methods for the remaining life of mechanical equipment are mainly divided into two categories: data-driven methods and physical model-based methods. The data-driven methods mainly include artificial intelligence algorithms[2]. The advantage is to avoid the establishment of complex actual models. The disadvantage is that a large amount of data support is required. In the literature [3], based on the fault tree analysis method, a transformer model was established, and the transformer running curve was fitted by least squares method to calculate the years of the transformer life and the average trouble-free working time.

The physical model-based method mainly predicts the life from the perspective of fatigue damage and aging mechanism[4]. The advantage is that there is less data to rely on. The disadvantage is that it is more difficult to establish an accurate actual model. The literature [5] is characterized by the average degree of polymerization of transformer insulating paper cellulose, and the chemical reaction kinetics is used to predict the residual life of transformer aging.

The variation of the electromagnetic force during the operation of the distribution transformer is complicated, and it is difficult to establish a physical model. Therefore, this paper proposes an
improved hidden semi-Markov model based on operational data to avoid the establishment of complex physical models. With less historical data, the remaining life can be predicted more accurately than the traditional model. Finally, the accuracy of the remaining life prediction of the improved model is verified by a practical example of a transformer.

2. Hidden semi-Markov model

Hidden semi-Markov model has good modeling and analysis capabilities for equipment fault prediction process. It consists of the following four basic elements:

1) The initial state probability distribution \( \pi = \{ \pi_i \} = P\{ s_1 = i \}, 1 \leq i \leq N \), where N is the number of states in the model and \( s_1 \) is the state of the system at the initial time.

2) The state transition probability distribution \( A = \{ a_{ij} \} \), where \( a_{ij} = P\{ s_{t+1} = j | s_t = i \} \). This matrix represents the transition probability between states.

3) The observation probability distribution \( B = \{ b_{ik} \} \), where \( b_{ik} = P\{ v_k | s_t = i \} \), \( 1 \leq k \leq M \), and M is the number of observations in state i. The matrix represents observations observed at different times to characterize the state.

4) The state duration matrix \( D = \{ P_i(d) \} \), \( 1 \leq i \leq N \). This matrix represents the probability of the event of staying in i for d time units.

It can be seen from the above analysis that the state transition matrix in HsMM defaults to a constant, and there is a large error between the actual remaining life of the remaining life prediction result. Therefore, the state transition coefficient is introduced according to the system performance degradation process, making the state transition matrix a time-varying matrix that changes with time, which is more in line with the actual performance degradation process of the system.

3. Degradation model of distribution transformer

A typical distribution transformer performance degradation curve is shown in Figure 1. It can be divided into three phases. The first stage is the smooth degradation stage, and the system performance changes remain basically unchanged. The second phase is the linear degradation phase, where the performance of the system changes evenly. The third stage is the accelerated degradation stage and the performance of the system deteriorates drastically.

Using the historical life data of the distribution transformer, the Baum-Welch algorithm is used to train the HsMM to obtain the state transition probability matrix A. However, when the system is in a state, the state transition probability of continuing to maintain the current state will decrease, and the transition probability to other states will increase. In addition, when the system is in different states, the change in state transition probability is inconsistent. Therefore, it is necessary to convert the fixed state transition matrix A into a time varying state transition matrix \( A(t) \) that changes with time. For the three degradation stages of the system, the time-varying state transition matrix is analyzed separately.

1) Smooth degradation period. At this stage, the state transition probability is constant over time, ie:
\[ a_s(t) - a_s(t + \Delta t) = \theta_1 \]  

where \( \theta_1 \) is a constant and \( \theta_1 \geq 0 \); \( \Delta t \) is a fixed interval between two observation times. Since \( \sum_{j=1}^{N} a_{ij} = 1 \), the variable \( \theta_1 \) needs to be assigned to \( a_{ij}(t + \Delta t) \). According to the hypothesis, the state transition probability of the system at the next observation time is:

\[ a_s(t + \Delta t) = a_s(t) - \theta_1 \]  

(2a)

\[ a_s(t + \Delta t) = a_s(t) + \frac{\theta a_s(t)}{\sum_{j=1}^{N} a_{ij}(t)} \]  

(2b)

According to the equation (2), the relationship between the state transition probability at the current time and the state transition probability when entering the state can be obtained:

\[ a_s(t_{k\Delta t}) = a_s(t = 0) - k\theta_1 \]  

(3a)

\[ a_s(t_{k\Delta t}) = a_s(t = 0) + \frac{k\theta a_s(t = 0)}{\sum_{j=1}^{N} a_{ij}(t = 0)} \]  

(3b)

2) Linear degradation period. At this stage, the state transition probability increases linearly with time, ie:

\[ \frac{a_s(t) - a_s(t + \Delta t)}{a_s(t)} = \theta_2 \]  

(4)

Similarly, the relational expression between the state transition probability at the current time and the state transition probability when entering the state can be obtained:

\[ a_s(t_{k\Delta t}) = (1 - \theta_2)^k a_s(t = 0) \]  

(5a)

\[ a_s(t_{k\Delta t}) = a_s(t = 0) + \frac{k\theta a_s(t = 0) a_s(t = 0)}{\sum_{j=1}^{N} a_{ij}(t = 0) \sum_{j=1}^{N} (1 - \theta_2)^{j-1}} \]  

(5b)

3) Accelerated degradation period. In this degenerate stage, the state transition probability changes over time in an exponential manner, ie:

\[ \frac{a_s(t) - a_s(t + \Delta t)}{a_s(t)} = a_s^{\theta}(t) \]  

(6)

The relation between the state transition probability at the current moment and the state transition probability when entering the state is:

\[ a_s(t_{k\Delta t}) = \left[ a_s(t = 0)^{1 - \theta} \right]^{t} \]  

(7a)

\[ a_s(t_{k\Delta t}) = a_s(t = 0) + \frac{a_s(t = 0) - a_s(t = 0)^{1 - \theta} a_s(t = 0)}{\sum_{j=1}^{N} a_{ij}(t = 0)} \]  

(7b)

The initial state transition probability matrix \( A_0 \) is obtained by training history data. Under actual circumstances, if the system is not repaired during operation, its performance will gradually degrade over time and will only shift to a worse health state, therefore, when \( 1 \leq i < j \leq N, a_{ij} = 0. \)

\[ A_0 = \begin{bmatrix} a_{11} & \ldots & a_{1N} \\ \vdots & \ddots & \vdots \\ 0 & \ldots & a_{NN} \end{bmatrix} \]  

(8)

Combining equation (3), (5), and (7) with equation (8) can obtain time-varying state transition probability matrices with different degradation stages.

4. Distribution transformer remaining lifetime prediction process
According to the Baum-Welch algorithm, the historical state-of-life data can be trained to obtain the initial state transition probability matrix $A_0$ and the mean and variance of the system in each state duration, and the duration obeys the Gaussian distribution. Define the forward variable:

$$\alpha(i,d)=\text{pr}[O_1,\ldots,O_t|(q_t,\tau_t)=(S_i,d)]$$

It indicates the probability that the observation sequence is $O_1,\ldots,O_t$ with the state at time $t$ is $S_i$ and the duration is $d$. The calculation equation is:

$$\alpha(i,d)=\pi b(O_1)p(d) \quad (10a)$$

$$\alpha(i,d)=\alpha_{i-1}(i,d+1)b_1(O_{t+1})+\sum_{j=1}^{N}a_{i-1}(j,d)a_j(O_t)b_j(O_1)p(d) \quad (10b)$$

Similarly, define the backward vector:

$$\beta(i,d)=\text{pr}[O_{t+1},\ldots,O_T|(q_t,\tau_t)=(S_i,d)]$$

It indicates the probability that the state at time $t$ is $S_i$ and the duration is $d$ under the condition that the observation sequence is $O_{t+1},\ldots,O_T$. The calculation equation is:

$$\beta_t(i,d)=1 \quad (12a)$$

$$\beta_t(i,1)=\sum_{j\neq i} a_{i,j}(O_{t+1})(\sum_{d=1}^{D} p(d)\beta_{t+1}(j,d)) \quad (12b)$$

$$\beta_t(i,d)=b(O_{t+1})\beta_{t+1}(i,d-1) \quad (12c)$$

Define variables:

$$\eta(i,d)=\text{pr}[O_{t+1},q_t=s_i,\tau_t=d]$$

Then the mean and variance of the state duration are:

$$\mu(i)=\frac{\sum_{t=1}^{T} \sum_{d=1}^{D} \eta(i,d)d}{\sum_{t=1}^{T} \sum_{d=1}^{D} \eta(i,d)} \quad (14a)$$

$$\sigma(i)=\frac{\sum_{t=1}^{T} \sum_{d=1}^{D} \eta(i,d)d^2}{\sum_{t=1}^{T} \sum_{d=1}^{D} \eta(i,d)}-[\mu(i)]^2 \quad (14b)$$

The state transition probability of the three degradation stages can be calculated by using the EM algorithm to calculate the value of the state transition coefficient based on the measurement data from the start to the current time. By comparing the probability that the system stays in the current state $a_{ij}(t)$ and the probability of transferring to other states $a_{ij}(t)$ ($1 \leq i \neq j \leq N$), when $a_{ij}(t) < a_{ij}(t)$, the system performance degradation state is considered to be transferred, from the current state $i$ to Other states $j$, thus using different state transition coefficients $O$ to calculate the time-varying state transition matrix. After selecting the state transition matrix, the values of the obtained state transition matrix are substituted into the Baum-Welch algorithm to find the corresponding mean and variance.

The mean and variance of each state duration are obtained by equation (14), and the dwell time of the device in each state is:

$$D(i)=\mu(i)+\rho \sigma^2(i) \quad (15a)$$

$$\alpha(i,d)=\text{pr}[O_1,\ldots,O_t|(q_t,\tau_t)=(S_i,d)] \quad (15b)$$

It can be seen from equation (14) that the state duration varies with the state transition matrix. The improved HSMM transforms the fixed state transition matrix in the traditional HSMM into a time-varying state transition matrix and is continuously updated according to the online monitoring data. The state transition probability, as the state transition probability changes, the system will also change in the duration of the current state, which can give a more accurate residual life prediction value:

$$RUL'_{i} = D_{i} + \sum_{j=1}^{N} D(j) \quad (16a)$$

$$D_{i}' = D(i)[1 - \frac{1}{\Pi_{j} d_{j}'}] \quad (16b)$$
$RUL_t$ is the remaining lifetime after the system runs $t$ time; $D(j)$ is the duration of the system in the $j$ state; $D_i^t$ is the dwell time in the state $i$ after the system runs $t$ time. The process of life prediction based on time-varying state transition HSMM is shown in Figure 2.

Fig. 2. Remaining lifetime prediction flow chart

5. Analysis of examples

In this paper, using the life-cycle operational data of a distribution transformer of Baoding and the vibration signal of the transformer as the original characteristic signal to determine the degradation state, the relative energy is extracted as a status characterization. The performance degradation curve of the distribution transformer conforms to the general trend of Figure 1, as shown in Figure 3. As the experiment progresses, the performance degradation curve of the distribution transformer can be roughly divided into three stages, which is consistent with the analysis in the previous section.

Fig. 3. Distribution transformer lifetime curve

According to the above life prediction process, the whole life history data is trained to obtain the initial state transition matrix of three states and the mean value and variance of each state duration, as shown in Table 1 and Table 2, respectively.

Table 1. The state transition matrix at beginning

| State      | Smooth | Linear | Accelerated |
|------------|--------|--------|-------------|
| Smooth     | 0.8848 | 0.0789 | 0.0363      |
| Linear     | 0      | 0.7989 | 0.2011      |
| Accelerated| 0      | 0      | 1           |

Table 2. The mean and variance of duration time at beginning

| State   | Mean/h | Variance/h |
|---------|--------|------------|
| Smooth  | 569    | 34.5       |
After the system runs for 200h, the mean and variance of the state transition probability and state duration of the system are re-estimated. At this point, the system is in a smooth degradation phase. Substituting it into equation (3), the state transition probability at this time is calculated, and the results as shown in Table 3 are obtained.

| State      | Smooth     | Linear | Accelerated |
|------------|------------|--------|-------------|
| Smooth     | 0.5254     | 0.4369 | 0.0377      |
| Linear     | 0.8375     | 0.1625 |             |
| Accelerated| 0          | 0      | 1           |

Substituting the values in Table 3 into equation (14), the mean and variance of each state duration is obtained, as shown in Table 4.

| State      | Mean/h | Variance/h |
|------------|--------|------------|
| Smooth     | 359    | 22.3       |
| Linear     | 505    | 30.7       |
| Accelerated| 46     | 2.3        |

Comparing Tables 1, 3 and Tables 2, 4, it can be found that the probability of the distribution transformer shifting from the stationary degradation state to the stationary degradation state gradually decreases with time, and the probability of shifting to other degradation states gradually increases. It is consistent with the actual situation.

In order to compare the accuracy of the time-varying state transfer HSMM and the traditional HSMM life prediction method, the relative error is selected as the evaluation index, and its expression is:

\[ E = \left| \frac{RUL_{\text{actual}} - RUL_{\text{predict}}}{RUL_{\text{actual}}} \right| \times 100\% \]  

(17)

Select 5 sets of distribution transformer with different working times, and use the method proposed in this paper to compare with the life prediction results obtained by the traditional HSMM life prediction method, as shown in Table 5.

| RUL/h | Predictive value/h | Relative error/% | Predictive value/h | Relative error/% |
|-------|--------------------|------------------|--------------------|------------------|
| 10    | 9.96               | 0.4              | 9.52               | 4.8              |
| 30    | 29.56              | 1.47             | 27.88              | 7.07             |
| 200   | 195.02             | 2.49             | 188.12             | 5.94             |
| 500   | 476.43             | 4.71             | 552.08             | 10.42            |
| 800   | 751.50             | 6.06             | 710.33             | 11.21            |

It can be seen from Table 5 that compared with the traditional HSMM method, the relative error of the method is small and the stability is within 5%. In the same degraded state, the accuracy of the proposed method gradually increases with time. This is because when more data is collected in a degraded state, the resulting remaining duration is more accurate.

6. Conclusion
In this paper, the state transition matrix is regarded as a fixed matrix in the traditional HSMM, which causes the residual life prediction to have a large error. It is proposed to use the time-varying state transition matrix to improve the accuracy of the remaining life prediction. Three different state
transition coefficients are proposed for different degradation stages of the system. The state transition coefficient is dynamically estimated according to the real-time monitoring data, the state transition probability matrix at different moments calculated, and the remaining life of the system in the current state updated in real time. Combined with the duration of the remaining health conditions given by the traditional HSMM, the accuracy of life prediction can be improved. Finally, the validity and accuracy of the method are verified by the life test of the distribution transformer.

7. References
[1] Yin J, Zhu Y and Yu G 2013 Transactions of China Electrotechnical Society. 28(1) 158-164
[2] Cheng F, Qu L, Qiao W and Hao L 2019 IEEE Transactions on Industrial Electronics. 66(6) 4738-4748
[3] Zhai B, Sun P and Ma J Power System Technology. 35(5) 127-131
[4] Jie M, Chen Z and Zhen Q Modern Manufacturing Engineering. (10) 148-152
[5] Wei J, Zhang G and Dong M High Voltage Engineering. 35(3) 544-550