Eng Int Syst (1998) 3: 141–146

Engineering
Intelligent Systems

Quality improvement of individual cores of distribution transformers using decision trees

N Hatziargyriou* P Georgilakis† D Spiliopoulos† and J Bakopoulos†

* Electrical Energy Systems Lab, National Technical University of Athens, Greece.
† Schneider Electric AE, Elvim Factory, Inofyta Viotia, Greece

The estimation of individual core losses of wound core distribution transformers (T/Fs) is very important because these constitute one of the main parameters of their quality and moreover core costs represent about 30% of the total T/F material cost. In addition, accurate calculations of individual core actual losses is extremely difficult, since actual losses show a divergence up to ±20%, in relation with the theoretical individual core losses. In this paper a decision tree (DT) method is presented, suitable for the prediction of individual core losses in wound core distribution transformers (T/Fs) at the early stages of the production process. For the creation of the learning and test set actual industrial measurements are used. Data comprise grain oriented steel electrical characteristics, and quality control measurements of cores production line. The resultant DT presents a success rate of 94 %. Based on this DT, rules comprising the most important parameters and their threshold values can be derived. These are used to improve the quality of individual cores by reducing their actual losses.

Keywords: Transformer Quality Improvement, Individual Core Losses, Decision Trees

1. INTRODUCTION

Accurate calculations of individual core losses are very difficult, due to the fact that many parameters, both qualitative and quantitative, influence the final product (individual core). These parameters are divided into three categories:

- Magnetic material parameters
  The most representative parameters of this category are:
  - Supplier
  - Thickness of magnetic material (0.23, 0.27, 0.30 mm)
  - Type of magnetic material (M3, M4, M5, Hi-B)
  - Specific losses of material
  - Resistivity of surface insulation
  - Hardness of material
- Constructional parameters
  The most representative parameters of this category are:
  - Rated induction
  - Thickness of core leg
  - Width of core leg
  - Height of core window
  - Width of core window
- Production procedure parameters
  The most representative parameters of this category are:
  - Annealing conditions
  - Destruction grade of insulation
  - Mechanical stresses during the formation of core
  - Real weight of core
  - Cutting quality of magnetic material

In order to study the impact of some of these parameters on the quality of individual core losses, the DTs method is used. The Decision Tree methodology, belonging to the Inductive Inference methods [1] has been further formulated in [2-3] . This method is recently applied to the solution of problems in power and industrial process systems [4-7]. In this paper the application of DTs for quality improvement of individual core of distribution transformers is presented.

The basic idea is to build decision rules in the form of a binary DT from a preclassified Learning Set consisting of Measurement Sets (MS) of cores. The Learning Set is based on actual industrial measurements. Each leaf of the DT carries the quality of individual core classified as acceptable or non-acceptable according to a given criterion concerning its losses and each interior node specifies an attribute to be tested with branches corresponding to the possible values of that attribute.

2. STRESS RELIEF ANNEALING

Important internal stresses are generated during the manufacturing process of wound cores and therefore a stress relief annealing is essential to restore the initial properties of the steel.

Stress relief annealing is designed to:

- reduce mechanical stresses in the steel to a minimum,
- prevent contamination of the steel with oxygen and/or carbon,
- and retain or enhance the insulation quality of the steel coating.

These qualities are generally achieved by annealing at temperatures in the 760 to 860 0C range, in a protective environment. The protective environment most widely used is pure, dry nitrogen which will protect the steel from oxidation. The environment may contain up to 2% hydrogen. Annealing in batch furnaces is especially recommended when annealing operations must be intermittent, or when the dimensions of the parts to be annealed are very different.

The annealing cycle followed in our application is divided into four stages:

- <u>Starting and heating up stage</u>
  The objective is to avoid oxidation and to normally achieve the temperature of 825 ºC. The duration of this cycle is between 2.75 and 3.25 hours. The nitrogen supply for the first hour is 14 m³/h. For the rest of the period a mixture of 98% N_2 and 2% H_2 is supplied at a 10 m³/h rate.
- <u>Soaking stage</u>
  The goal is that all cores in the load must have homogeneous temperature distribution. The duration of this phase is 2.5 hours, at 825 ºC. A mixture of 98% N_2 and 2% H_2 is supplied at a 8 m³/h rate.
- <u>Slow cooling stage</u>
  The target is to cool the load slowly, in order to avoid the development of internal stresses in the cores. The duration of this cycle is 2 hours. The final temperature is 700 ºC. A mixture of 98% N_2 and 2% H_2 is supplied at a 8 m³/h rate.
- <u>Fast cooling stage</u>
  The objective is to reduce the temperature to 380 ºC, in order to avoid oxidation of cores, when they are going to be exposed to the environment. The duration of this cycle is 3.5 hours. A mixture of 98% N_2 and 2% H_2 is supplied at a 8 m³/h rate, until the temperature is higher or equal to 600 ºC. When the temperature is 600 ºC, the supply of N_2 is 14 m³/h.

The total duration of the cycle must not exceed 11.5 hours. The reason is that in a 24-hours period, it is desirable two annealing cycles to be implemented.

3. DESIGN OF WOUND CORE T/FS

The thickness, E_u (in mm), of core leg is calculated from [7]:

$$E_u = \frac{VPT * 34.9 * 10^5 * 25.4^2}{2 * CSF * D * B * f}$$

VPT: T/F volts per turn
CSF: Core space factor (typical value: 0.965)
B: rated induction (in Gauss)
D: width of core leg (in mm)
f: rated frequency.

In order to construct a 3-phase distribution transformer, 2 small cores (with width of core window equal to F1), and 2 large cores (with width of core window equal to F2) must be assembled. The core constructional parameters are shown in Figure 1.

The following relation between the width of core window of the large and small core exists:

$$F2 = 2 * F1$$

The mean turn, CMT1 (in mm), of the small core is:

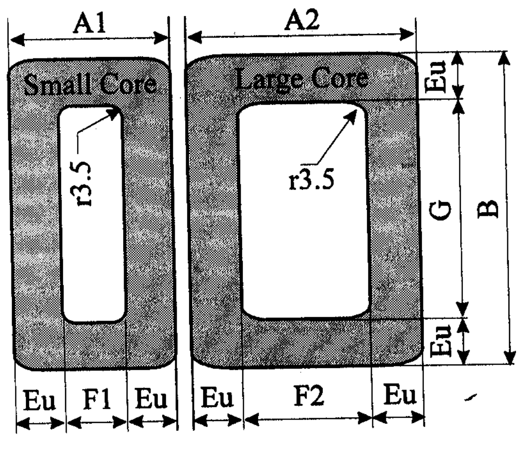

Figure 1 Core constructional parameters

$$CMT1 = 2 * (F1 + G) + 2 * 3.14 * \left(\frac{E_u}{2} + 3.5\right) - 8 * 3.5$$

The theoretical weight, CTW1 (in Kg), of the small core is:

$$CTW1 = CMT1 * D * E_u * CSF * g_{MS} * 10^{-6}$$

where g_{MS} (in g/cm^3) is the magnetic steel density.

The mean turn, CMT2 (in mm), of the large core is:

$$CMT2 = 2 * (F2 + G) + 2 * 3.14 * \left(\frac{E_u}{2} + 3.5\right) - 8 * 3.5$$

The theoretical weight, CTW2 (in Kg), of the large core is:

$$CTW2 = CMT2 * D * E_u * CSF * g_{MS} * 10^{-6}$$

The theoretical total weight, CTW, of the T/F cores is:

$$CTW = 2 * (CTW1 + CTW2)$$

The theoretical individual core losses, W1 (in Watt), of the small core are obtained from:

$$W1 = WPK_1 * CTW1$$

Where WPK1 are the theoretical core specific losses at the rated induction (see Figure 2).

The theoretical individual core losses, W2 (in Watt), of the large core are:

$$W2 = WPK1 * CTW2$$

4. OVERVIEW OF DT METHODOLOGY

The Decision Tree methodology [3] is a non-parametric learning technique able to produce classifiers about a given problem in order to reduce information for new, unobserved cases. The DT is a tree structured upside down, built on the basis of a Learning Set (LS). The LS comprises a number of preclassified states defined by a list of candidate attributes.

In order to detect if a node is terminal, i.e. "sufficiently" class pure, the classification entropy of the node with a minimum preset value Hmin is compared. If it is lower than H_{min}, then the node is sufficiently class-pure and it is not further split. Such nodes are labeled LEAVES. Otherwise, a suitable test is sought to divide the node, by applying the optimal splitting rule. In the case that no test can be found with a statistically significant information gain, the node is declared a DEADEND and it is not split.

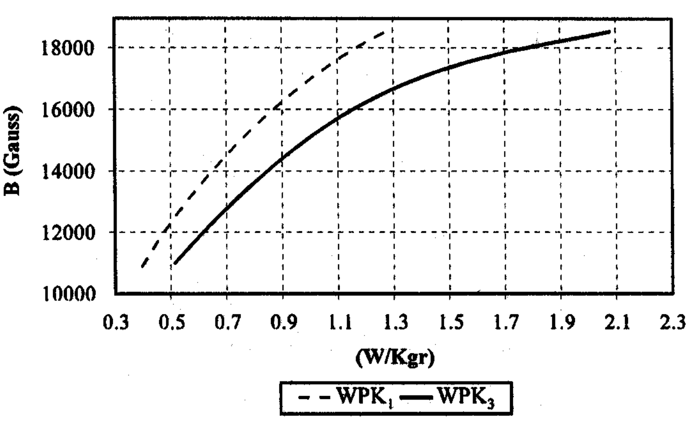

Figure 2 typical losses curve

DTs are tested using test sets (TS), comprising a number of similar, preclassified, but independent MS. The class of each of these MS is compared to the class of the terminal node finally led to by applying the tests of the various non-terminal nodes. This comparison provides the DT classification error rate.

Each node possesses a subset of MS with the following characteristics:

E_n: the MS subset of node n of the DT.
N: size (number of MS) of E_n.
n_a: number of acceptable MS in E_n.
n_{na}: number of non-acceptable MS in E_n.

The relative frequencies of acceptable and non-acceptable MS for node n will be:

$$f_a = \frac{n_a}{n_a + n_{na}} = \frac{n_a}{N} \quad \text{and} \quad f_{na} = \frac{n_{na}}{n_a + n_{na}} = \frac{n_{na}}{N}$$

The entropy of E_n with respect to the class partition of its elements, is defined as:

$$H_c(E_n) = -(f_\alpha \log f_\alpha + f_{n\alpha} \log f_{n\alpha})$$

A test T is defined at node n as:

$$T : A_i \leq t$$

where A_i is the value of attribute i of a particular MS and t is a threshold value. By applying the test T to all MS of node n, E_n is split into two subsets E_{n1} and E_{n2}:

$$E_{n1} = \{MS \in E_n : A_i \leq t\} \text{ and } E_{n2} = \{MS \in E_n : A_i > t\}$$

If n_i : number of MS in E_{ni}, i=1,2 the corresponding frequencies are given by:

$$f_1 = \frac{n_1}{n_1 + n_2} = \frac{n_1}{N} \quad \text{and} \quad f_2 = \frac{n_2}{n_1 + n_2} = \frac{n_2}{N}$$

The entropy of E_n with respect to the partition induced by T is:

$$H_T(E_n) = -(f_1 \log f_1 + f_2 \log f_2)$$

$H_T(E_n)$ is a measure of the uncertainty of the outcome of test T. The mean conditional entropy of E_n, given the outcome of test, corresponds to the residual entropy after the application of T and is defined as:

$$H_c(E_n \mid T) = f_1 H_c(E_{n1}) + f_2 H_c(E_{n2})$$

The information gained from the application of test T is expressed by the achieved reduction of the learning subset entropy:

$$I(E_n ; T) = H_c(E_n) - H_c(E_n \mid T)$$

A more objective (less biased) estimator of the merit of test T is provided by the normalized information gain, defined as:

$$C(E_n ; T) = \frac{2\, I(E_n;T)}{H_c(E_n) + H_T(E_n)} \in [0,1]$$

Under the hypothesis of no correlation between the test T and the class partition in the Universe U of the MS (resp. U_n), that is for zero actual increase in information, the random variable $N*I(E_n;T)$, which is an estimator of the total actual information gain, is X^2-distributed with 1 degree of freedom and its expected value is positive and inversely proportional to the size of the subset E_n.

$$N * I(E_n ; T) \sim X^2(1)$$

If a is the risk level of not detecting situations of only apparent information gain and X_{cr} the value such that $P(X>X_{cr})= a$, where X a random variable following an X^2 distribution with 1 degree of freedom, then the following statistical test can be formulated: "The node splitting test T is rejected as uncorrelated with the class partition if $Q1=N*I(E_n;T)<X_{cr}$", where N is the number of learning states in E_n.

5. STATISTICAL DESIGN OF EXPERIMENTS

In order to study the annealing cycle and to investigate its impact on the individual core losses of wound core distribution transformers, eight tests are designed.

In every one of the eight tests the following six parameters are varied:

- Annealing final temperature (ATTR1).
  Values : 825 0C (L - Low), and 855 0C (H - High)
- Temperature rising time (ATTR2).
  Values : 3 hours (L), and 4 hours (H)
- Furnace opening temperature (ATTR3).
  Values : 250 0C (L), and 350 0C (H)
- Duration of constant temperature (ATTR4).
- Position of core in the furnace (ATTR5).
  Values : Down (L), and Up (H)
- Protective atmosphere (ATTR6).
  Values : 100% N2 (L), and mixture of 98% N2 and 2% H2 (H)

Due to production reasons, all the experiments must follow the constraint that the total duration of the annealing cycle must not exceed 11.5 hours. In order to achieve that, the duration of slow and fast cooling stages, which are not included to the previous six attributes, are appropriately selected.

The parameters characterizing each of the 8 tests are shown in Table 1.

Table 1 Conditions of the various annealing tests.

	ANNEALING TEST Nb							
	1	2	3	4	5	6	7	8
ATTR1	L	H	H	L	L	H	H	L
ATTR2	L	L	H	H	H	H	L	L
ATTR3	L	H	H	L	H	L	L	H
ATTR4	L	L	H	H	L	L	H	H
ATTR5	H	L	H	L	H	L	H	L
ATTR6	H	H	H	H	L	L	L	L

In order to take in account all the combinations of the six parameters with two values (Low, and High), 32 experiments are required. Time and finance constraint reasons however lead us to do 8 experiments. It can be seen that 4 experiments are carried out with the low value of each attribute, and the other 4 with the high value.

Two additional variables are selected as candidate attributes:

- Actual over theoretical core weight ratio (ATTR7).

$$ATTR\ 7 = \frac{ActualweightOfCore}{TheoreticalWeight\ OfCore}$$

- Specific losses of core magnetic material (ATTR8).

$$ATTR\ 8 = MaterialLossesOf\ Core$$

where MaterialLossesOfCore: specific losses (Watt/Kg at 15000 Gauss) of core magnetic material.

All tests were done using the same 160 kVA T/F design, and also the same supplier of cores magnetic material. The magnetic material was M3, 0.23 mm. For every one of the eight tests, 96 (48 small, and 48 large) cores were constructed. It must be noticed that all cores were annealed at the same batch furnace.

6. APPLICATION

For the creation of the learning and test set actual industrial measurements are used. 768 Measurement Sets are collected, for this purpose. The 3/4 of the MS are used as learning set, and the rest as test set.

The criterion for the classification of individual core losses is the following: one core is NON-ACCEPTABLE if the actual specific core losses (measured by the Quality Control Department) are greater than 15% of the theoretical specific core losses (calculated by T/Fs Design Department). Otherwise, it is labeled ACCEPTABLE. In Table 2 the percentage of acceptable cores per annealing test, is shown.

In Figure 3 a characteristic DT is illustrated, developed with the 8 attribute list and 0.999 confidence level. Its success rate, tested with the TS comprising 192 MS is 94%. The notation used for the DT nodes is explained in Figure 4.

The Safety Index of a node is defined as the ratio of the acceptable MS in the subset E_n of node n to the total num-

Table 2 Percentage of acceptable cores per test

	ANNEALING TEST Nb							
	1	2	3	4	5	6	7	8
Per (%)	94	95	93	69	94	98	98	93

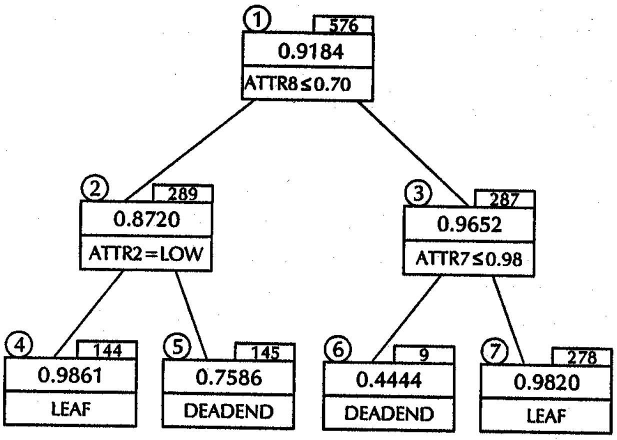

Figure 3 DT developed using the 8 attributes set

ber of MS in E_n. If the safety index of a terminal node is greater than 0.5, then the MS "falling" to this node are characterized as ACCEPTABLE, otherwise as NON-ACCEPTABLE.

Increasing the confidence level, the size of the developed DTs decreases and vice versa. Generally, the calculated accuracy of the DTs, as expected by the percentage of successful classifications, is proportional to their size, i.e. there is a trade-off between the size of the DTs and their reliability in classifying the individual core losses.

Using the 192 MS test set, the classifying error rate of each particular terminal node can be determined. It has been found that deadends are not always characterized by low classifying accuracy. For example, in the DT of Figure 3, deadend 5 is 87% successful in classifying 54 MS, due to the fact that this node possess sizeable MS subsets of 145 MS, while its safety index is 0.7586 (i.e. its impurity is low, but not low enough to be declared leaf). Hence, although it is desirable for the DTs to comprise as many leaves as possible, the number of leaves/deadends is not a direct measure of the DT quality.

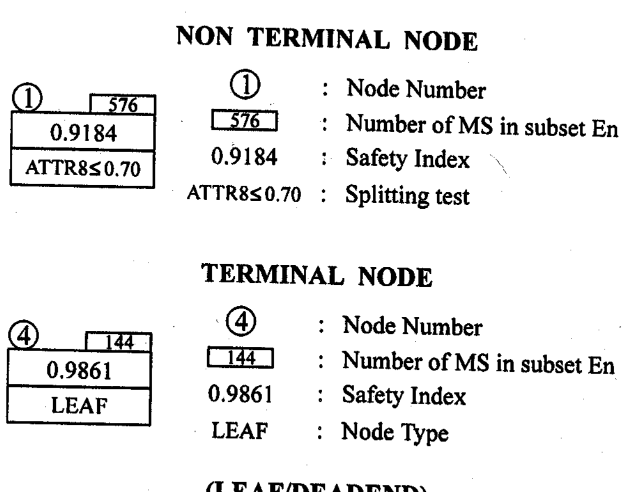

Figure 4 Notation of the DT's nodes

7. PRACTICAL USE OF RESULTS

The 3 attributes appearing in the node splitting tests of the DT of Figure 3 in decreasing order of significance are: ATTR8, ATTR2, and ATTR7. Parameter ATTR8 reflects the quality of the material, as it is equal to the specific losses (Watt/Kg at 15000 Gauss) of core magnetic material. Parameter ATTR2 represents the temperature rising time of the annealing cycle, while parameter ATTR7 expresses the actual over theoretical core weight ratio.

The selection of these attributes is reasonable and expected, since they are all related to the quality of individual core. It is notable that the only variable, relevant to the annealing cycle, that appears in the node splitting tests of the DT is ATTR2. This is due to the fact that ATTR2, ATTR4, and also the duration of the slow and fast cooling stages are strongly correlated, since the annealing time is constant. On the other hand ATTR5, which declares the position of core in the furnace, is not important.

Based on the Decision Tree of Figure 3, rules useful for the Transformers Production Department are derived. It is desirable to construct cores leading to nodes 4 and 7, if it is technoeconomically feasible. These nodes have safety index greater than 98%.

The measurement sets following the rule ATTR8>0.7 and ATTR7≤0.98 are lead to node 6, and characterized as non-acceptable. In order to avoid this, the T/Fs Production Department must increase ATTR7. This is equivalent to increasing the real weight of core by adding more magnetic material, so that the actual over theoretical core weight ratio (ATTR7) is greater than 0.98.

Given the quality of the magnetic material (ATTR8) the most suitable annealing test can be selected as follows:

- if ATTR8≤0.7 the annealing test 7 must be selected. The reason is that in this case it is desirable to lead to node 4. From node 2 splitting rule it can be derived that ATTR2 must be Low. Additionally, from Table 2 it can be seen that the best annealing cycles are those of tests 6 and 7, which lead to 98% acceptable cores. From these two tests, only test 7 has the ATTR2 equal to Low (see Table 1). At a first impression this result seems unexpected. However, this can be easily explained if the total annealing cycle is considered. This includes not only the temperature rising time, but also the duration of constant temperature and the slow and fast cooling stages.
- if ATTR8>0.7 the standard annealing cycle must be selected. The reason is that node 3 splitting rule is independent from the annealing variables.

8. CONCLUSIONS

In this paper, Decision Trees are applied for the quality improvement of individual core losses of wound core distribution transformers. The basic steps in the application of the method to cores of a 160 kVA transformer design, like the generation of the learning set and test set, the selection of candidate attributes and the derivation of a characteristic decision tree are presented.

With the LS and TS used and for the selected candidate attribute sets, the DT method achieve a total classification success rate of 94 %, which renders it very suitable for quality improvement of individual core of distribution transformers.

ACKNOWLEDGEMENTS

The authors would like to thank the General Secretariat of Research and Technology of Greece for financing the research programme entitled: "Development of a Quality Control System of Distribution Transformer Magnetic Cores Using Artificial Intelligence Techniques" within the YPER '94 Research Programme. Help from Schneider Electric AE operators in collecting industrial measurements is gratefully acknowledged.

REFERENCES

[1] Quinlan J.R, "Induction of Decision Trees", Machine Learning, Vol. 1, No. 1, pp. 81-106, 1986.

[2] Wehenkel L., Cutsem Th. Van, Ribbens-Pavella M., "An Artificial Intelligence Framework for On-Line Transient Stability Assessment of Power Systems", IEEE Trans on Power Systems, Vol. 4, No.2, pp. 789-800, May 1989.

[3] Wehenkel L., Pavella M., "Decision Trees and Transient Stability of Electric Power Systems", Automatica, Vol. 27, No.1, pp. 115-134, 1991.

[4] Hatziargyriou N.D., Contaxis G.C., Sideris N.C., "A Decision Tree Method for On-Line Steady state Security Assessment", IEEE, paper 93 SM 527-2 PWRD.

[5] Hatziargyriou N.D., Papathanassiou S.A, Pecas-Lopes J.A, Acker V. Van, "Pattern Recognition versus Decision Trees Methods – A Case Study in Fast Dynamic Security Assessment of Autonomous Power Systems with a Large Penetration From Renewables", ISAP, Montpellier, France, 1994.

[6] Pecas Lopes J.A., Hatziargyriou N.D, "State of the art about on-line security assessment of electric power systems", IV Symposium of Specialists in Electric Operational and Expansion Planning, Foz do Iguacu, Brazil, May 1994.

[7] Georgilakis P.S., Bakopoulos J.A., Hatziargyriou N.D, "A Decision Tree Method for Prediction of Distribution Transformer Iron Losses", to be presented to 32nd UPEC, Manchester, September 1997.