

데이터마이닝 기법을 이용한 주상변압기 고장유형 분석 및 복구 예측모델 구축에 관한 연구

論 文

57-9-4

Fault Pattern Analysis and Restoration Prediction Model Construction of Pole Transformer Using Data Mining Techniques

黃友炫* ·金子姬[†] ·蔣完成** ·洪淳植*** ·韓得洙[§]

(Woohyun Hwang · Ja-Hee Kim · Wan-Sung Jang · Jung-Sik Hong · Deuk-Su Han)

Abstract - It is essential for electric power companies to have a quick restoration system of the faulted pole transformers which occupy most of transformers to supply stable electricity. However, it takes too much time to restore it when a transformer is out of order suddenly because we now count on operator in investigating causes of failure and making decision of recovery methods. This paper presents the concept of 'Fault pattern analysis and Restoration prediction model using Data mining techniques', which is based on accumulated fault record of pole transformers in the past. For this, it also suggests external and internal causes of fault which influence the fault pattern of pole transformers. It is expected that we can reduce not only defects in manufacturing procedure by upgrading quality but also the time of predicting fault patterns and recovering when faults occur by using the result.

Key Words : Pole Transformer, Data mining, Decision Tree, CART, Pole Transformer Fault Pattern, Fault Restoration Prediction Model

1. 서 론

주상변압기는 우리나라에 전기가 처음 도입된 1887년 이래 지속적으로 증가해 2006년 말 현재 전국 전력공급 변압기의 약 80%에 이르고 있다[1]. 주상변압기는 22,900V의 배전선로에 연결되어 220V 또는 380V의 전압으로 낮추어 대부분의 주택, 소규모 상가, 공장 등에 전력을 공급하는 전기기기이다[2]. 주상변압기는 열화로 인한 고장을 예방하고자 설치 후 13년이 경과되면 교체하고 있다. 그러나 최근의 고장현황을 분석한 결과 13년 미만의 주상변압기에서도 고장이 발생하고 있어 사용자에게 불편을 끼치고 있다[3]. 따라서 주상변압기 경과연수 외의 다른 고장요인에 대해서도 규명하는 것이 필요하다. 또한 규명된 고장요인을 바탕으로 고장유형 분석과 신속한 고장복구 방안에 관한 연구가 필요한 실정이다.

주상변압기의 고장을 감소시키고자 진행된 연구를 살펴보면 국내에서는 분광광도법이나 전기분석법을 이용하여 절연유 내의 구리와 푸르푸랄의 증가량이나 유중 가스를 토대로 주상변압기의 열화진단에 관한 연구가 진행된 바 있다[4],[5]. 이 연구는 변압기 내부의 절연유를 분석하여 고장을 예측하는 방법으로 발표되었으나 활선상태에서 변압기 절연유 채취가 곤

란하여 활용이 불투명하다. 또한 주상변압기의 2차 전류, 누설 전류, 외함온도 등을 원적으로 측정하여 열화를 진단하는 사례[6]의 경우 투자비가 많이 소요되어 실제 적용은 어려운 것으로 보인다. 이러한 연구사례는 변압기 외부환경요인에 대한 연관성은 규명되지 않았다. 국외 연구사례도 전력용 변압기나 전력계통에서 수집된 데이터를 인공신경망을 이용하여 고장을 예측하고자 하는 연구 등이 있으나[7][8], 중간과정이 상세히 표현되지 않아 관계성이 규명되지 않았고, 무선통신망을 이용한 변압기 열화진단 장치를 개발하여 2차측 전압변동과 전압과 전류파형을 측정하여 절연유의 열화판단 기준을 제시한 연구가 있으나[9], 고장유형이 입력변수로 설정되어 있어, 고장유형 예측에는 미흡하다고 하겠다. 데이터마이닝 기법을 이용한 고장유형 분석과 고장예측 그리고 고장발생시 고장복구 모델구축에 관한 연구는 아직 초기 단계에 있다. 따라서 선행 연구에서 미흡한 점을 과거 3년간 축적된 주상변압기 고장발생 자료를 토대로 데이터마이닝 기법의 의사결정나무를 이용하여 다음과 같이 4단계로 나누어 연구를 진행한다. 첫째는 여러 가지 데이터마이닝 기법 중 주상변압기 고장유형 분석에 적합한 데이터마이닝 기법과 알고리즘을 선택한다. 본 연구에서는 고장유형 분석에 적합한 의사결정나무의 CART 알고리즘을 선택하였다. 둘째는 선택한 데이터마이닝 기법을 적용하기 위한 주상변압기 고장에 영향을 주는 요인들을 선정하고 분석방법을 구체화 한다. 셋째는 주상변압기의 고장유형과 고장요인과의 관계성을 분석하기 위해 자료 수집 시 주상변압기 고장발생 일시, 설치된 장소, 고장 변압기 종류, 공급전압, 날씨, 온도, 습도 그리고 변압기 이용률 등과 같은 내용을 포함 한다. 넷째는 수집된 자료를 활용하여 변수 변환 표를 만들고 마스터데이터를 작성하여 자료변환과 정제를 거쳐 상용 데이터마이닝 도구인 SPSS사의 의사결정나무를 이용하여 고장유형을 분석한 다음 전개된 나무결

* 正 會 員 : 국립서울산업대학교 산업정보시스템전공 박사과정

[†] 교신저자, 正 會 員 : 국립서울산업대학교 정보시스템공학과
助敎授 · 工博

E-mail : jahee@snut.ac.kr

** 非 會 員 : 한국전력공사 배전계획처장

*** 非 會 員 : 국립서울산업대학교 정보시스템공학과 敎授 · 工博

§ 正 會 員 : 한국전력공사 과장대리

接受日字 : 2008年 7月 25日

最終完了 : 2008年 8月 12日

과를 토대로 고장발생 시 고장유형에 따라 신속히 고장을 복구하기 위한 모델을 제시한다.

이를 위한 본 논문의 구성은 다음과 같다. 제 2장에서는 주상변압기의 개요, 주요 고장요인과 현황을 분석하고 제 3장에서는 데이터마이닝 기법의 의사결정나무와 CART 알고리즘에 대한 이론적 배경을 검토한 다음 제 4장에서는 의사결정나무 구동에 필요한 자료를 종합하고 마스터테이블의 작성과정 및 나무생성 결과를 분석하여 제 5장에서 연구결과의 결론을 정리하였다.

2. 주상변압기의 개요 및 고장복구 절차

이 절에서는 주상변압기의 구조와 종류 그리고 어떠한 요인이 주상변압기의 고장 발생에 영향을 미치고 있는지를 알아본다. 또한 고장이 발생하는 요인 중 외부 환경요인과 내부 열화요인의 관계성을 도식화하였으며 최근 3년간 고장발생 현황 분석과 고장복구 절차에 대해 알아본다.

2.1 전력계통과 주상변압기

주상변압기는 22,900V의 고전압을 사용자가 편리하게 사용할 수 있도록 전압을 220V 또는 380V로 저하시켜주는 전기기기이다. 주상변압기의 구조는 크게 외함파고 고압 부싱, 저압 부싱으로 구성되어 있다. 주상변압기 내부에는 탭 절환기와 철심 그리고 코일로 구성되어 있으며, 고전압이 흐를 때 외함파고 코일 간 절연내력을 높이고 열 방산을 위해 절연유를 탭 절환기 부위까지 채우고 있다. 주상변압기는 1대로는 일반 가정용이나 소규모 상가에 공급하고, 3대를 조합할 경우에는 소규모 상가나 공장과 같은 사용자에게 전력을 공급하는데 사용하고 있다.

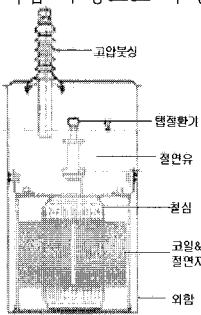


그림 1 주상 변압기의 구조
Fig. 1 Structure of Pole Transformer

현재 가장 많이 사용되고 있는 주상변압기는 22,900V 일 단 접지형으로 철심의 종류와 보유 기능에 따라 일반형, 자기진단형, 저손실형, 보호장치 내장형, 아몰퍼스형 등 5가지가 있으며 실제 사용되고 있는 형태는 그림 2와 같다.

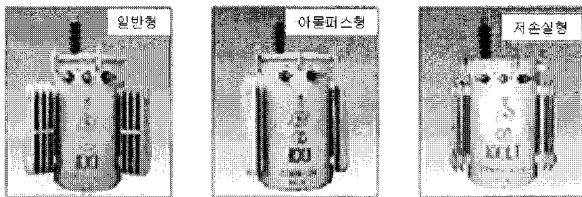


그림 2 주상 변압기의 종류
Fig. 2 Type of Pole Transformer

주상변압기는 제작당시 정격용량 100kVA를 기준으로 하여 설계기준상 표고 1,000m 이하의 곳에서 여름철에는 40°C의 고온과 겨울철에는 영하 20°C의 저온에서도 견디어야 한다.

주상변압기의 점검은 주상점검, 입고점검, 교체점검 및 설치 전 점검이 있다. 주상점검은 주상설치상태에서 연 1회 정기적으로 시행하고, 설치 후 매 1년마다 정기점검을 시행하며, 교체점검은 제작 후 13년이 초과된 변압기에 대해 실시한다. 입고점검은 각종 공사로 철거되는 주상변압기를 창고에 입고하기 전 시행하는 점검이고, 설치 전 점검은 사용자의 요청에 의해 주상변압기를 전주에 설치하기 전에 현장에서 시행하는 점검을 말한다.

2.2 주상변압기 고장에 영향을 주는 주요 요인

매년 통상적으로 발생하고 있는 주상변압기 고장유형에 대한 주요 요인은 그림 3과 같이 표시할 수 있다. 날씨는 맑음과 비 또는 눈이 오는 날에 따라 영향을 받으며 특히 낙뢰에 의한 고장발생 빈도가 높다. 계절의 경우 냉방부하가 많이 걸리는 여름철에 영향이 크고 과부하는 사용부하의 종류에 따라 130%를 초과하여 장시간 사용할 경우 고장의 요인이 된다. 또한 온도와 습도가 많을 경우, 제작당시 부품과 조립 시 작업자의 실수에 의해 고장이 발생하기도 하고 설치 후 사용연수가 오래 될수록 큰 영향을 받는다. 이러한 환경요인과 열화요인을 분류하면 다음과 같다.

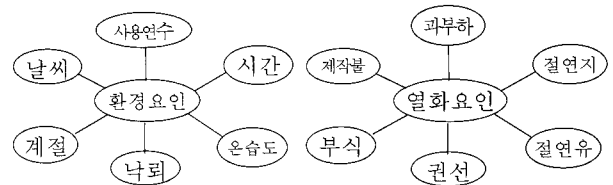


그림 3 주상변압기 고장 요인
Fig. 3 Fault Cause of Pole Transformer

2.3 주상변압기 고장 발생 현황

사용자들에게 전력을 공급하기 위해 사용되고 있는 주상변압기의 최근 3년간 고장발생 현황을 유형별로 살펴보면 표 1과 같다.

표 1 주상변압기 고장 현황
Table 1 Fault Status of Pole Transformer

구분	'04	'05	'06	
설치수량(만대)	165	173	178	
고장수량(대)	8,089	9,157	9,004	
고장 유형	자연열화	3,771	4,220	4,285
	자연현상	1,523	2,319	2,217
	과부하	1,129	905	811
	제작불량	431	415	446
	부식	203	263	322
	기타	1,032	1,035	923

표 1에서 보는 바와 같이 주상변압기의 고장률은 전체 설치수량과 비교해 보면 미미한 규모이나 고장유형은 자연 열화, 자연현상, 과부하, 제작불량 순으로 나타나고 있다.

2.4 주상변압기 고장 복구절차

주상변압기의 고장이 발생하면 사용자의 신고에 의해 고장위치를 파악하고 운영센터의 근무자가 현장에 이동하여 고장요인을 분석하고 복구처리하게 된다.

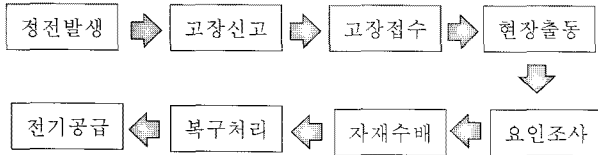


그림 4 고장복구 절차
Fig. 4 Fault Restoration Procedure

이 과정에서 고장접수 후 고장변압기의 교체가 완료되기 까지 소요되는 평균 시간은 현장까지 이동시간과 요인조사 그리고 교체용 변압기수배 등에 대부분의 시간이 걸리고 있다.

3. 데이터마이닝 기법과 CART 알고리즘

3.1 데이터마이닝의 정의와 적용사례

데이터마이닝이란 대용량의 데이터를 수집하여 유형별로 분류하고 데이터별 연관성을 분석하여 정보로 환원시키는 일련의 과정이라고 할 수 있다. 데이터마이닝을 이용하여 연구된 사례들을 살펴보면 신용카드 고객이탈을 방지하기 위한 가입고객의 특성분석, 통신회사 고객의 서비스 가입해 지를 예측하기 위한 모형의 체계화, 의사결정나무를 이용한 대학생의 전공 이탈자 예측과 2단 신경망을 이용한 원자력 발전소의 고장진단 등에 사용한 사례가 있고, 해외에서도 배전선로의 고장유형 분류와 신경망을 이용한 배전용 변압기의 진단에 사용된 경우는 있으나 데이터마이닝 기법을 전력기기 고장유형 분석과 고장복구 모델 구축에 적용하여 연구된 사례는 확인되지 않았다.

3.2 의사결정나무(Decision Tree) CART 알고리즘

일반적으로 가장 많이 사용하고 있는 데이터마이닝 기법 중 의사결정나무에 대한 알고리즘에 대해 살펴본다.

의사결정나무는 대량의 자료를 수집하여 의사결정을 하고 자하는 규칙을 나무구조로 도표화하여 목표 집단을 몇 개의 소집단으로 분류하거나 예측하는데 사용하는 방법이다. 자료의 분석과정이 나무모형으로 일목요연하게 표현되기 때문에 분류 또는 예측을 목적으로 하는 회귀분석이나 인공신경망이 판별분석보다 분석과정을 쉽게 나타낼 수 있다.

의사결정나무의 대표적인 알고리즘은 CHAID, CART, C4.5, QUEST 등이 있으며, SPSS, SAS와 같은 소프트웨어 회사들에 의해 상용화되어 있다. SPSS사의 Answer Tree에서는 CHAID, CART, QUEST 등 다양한 알고리즘을 이용하여 의사결정 나무분석을 수행할 수 있으며 표 2는 알고리즘을 비교한 것이다. 본 논문에서는 주상변압기의 고장유형을 분석하는데 사용된 자료가 명목형, 순서형, 연속형의 목표변수를 모두 구현이 가능하고 분리개수는 Binary로 전개되어 CHAID와 QUEST에 비해 정확성이 높을 것으로 판단되어 CART 알고리즘을 적용하였다.

표 2 데이터마이닝 알고리즘 비교

Table 2 Data Mining Algorithm

구분	CHAID	CART	QUEST
목표변수	명목, 순서, 연속	명목, 순서, 연속	명목
예측변수	명목, 순서, 연속	명목, 순서, 연속	명목, 순서, 연속
분리기준	카이제곱-검정	지니지수	카이제곱-검정
분리개수	Multiway	Binary	Binary

CART(Classification and Regression Trees) 알고리즘은 지니지수(Gini index) 또는 분산(Variance)의 감소량을 분리 기준으로 사용하여 이진분리(Binary Split, 부모마디에서 자식마디가 2개만 형성되게 함)를 수행하는 것을 말한다.

CART는 임의의 한 개체가 목표변수의 i번째 범주로부터 추출되었고, 그 개체는 목표변수의 j번째 범주에 속하며 오분류(Misclassification)될 확률은 P(i)P(j)가 된다. 여기서 P(i)는 각 마디에서 한 개체가 목표변수의 i번째 범주에 속할 확률이다. 또한 지니지수는 각 마디에서 불순도(Impurity) 또는 다양도(Diversity)를 재는 척도 중의 하나로

$$G = \sum_{j=1}^c P(j)(1-P(j)) = 1 - \sum_{j=1}^c P(j)^2 = 1 - \sum_{i=1}^c \left(\frac{n_i}{n}\right)^2 \quad (1)$$

와 같이 표현될 수 있다. n은 관찰치 수치고 n_i는 목표변수의 i번째 관찰치이다.

CART는 이 지니지수를 가장 잘 감소시켜주는 예측변수와 그 변수의 최적분리를 자식마디로 선택하는데 지니 지수의 감소량은 다음과 같이 계산된다.

$$\Delta G = G - \frac{n_L}{n} G_L - \frac{n_R}{n} G_R \quad (2)$$

여기서 n은 부모마디의 관찰치 수를 말하고 n_R과 n_L은 각각의 자식마디에서 관찰치 수를 의미한다. 즉, 자식마디로 분리되었을 때 불순도가 가장 작도록 자식마디를 형성하는 것이며 자식마디에서 불순도의 가중 합을 다음과 같이 최소화하는 것과 동일하다.

$$P(L)G_L + P(R)G_R = \frac{n_L}{n} G_L + \frac{n_R}{n} G_R \quad (3)$$

의사결정나무의 장점은 나무구조모형으로 표현이 되기 때문에 모형 사용자가 쉽게 이해할 수 있으며, 나무구조로부터 어떤 예측변수가 목표변수를 설명하기 위해서 더 중요한 것 인지를 쉽게 파악할 수 있다. 또한 두개이상의 변수가 결합하여 목표변수에 영향을 주는 관계성을 쉽게 알 수 있으며, 선형성(Linearity)이나 정규성(Normality) 또는 등분산성(Equalvariance) 등의 가정을 필요로 하지 않는 비모수적인 방법이다. 한편, 연속성 변수를 비연속적인 값으로 취급하기 때문에 분리의 경계점 부근에서는 예측오류가 클 가능성과 선형 또는 주 효과 모형에서와 같은 결과를 얻을 수 없는 한계점이 있으며 분석용 자료(Training Data)에만 의존하는 의사결정나무는 새로운 자료의 예측에서는 불안정할 가능성이 높다는 것이 단점이기 때문에 검증용 자료(Test Data)에 의한 교차타당성(Cross Validation) 평가나 가지치기에 의해서 안정성 있는 의사결정나무를 얻는 것이 바람직하다[10-15].

4. 단계별 데이터마이닝 절차와 분석 결과

본 논문에서는 데이터마이닝 기법의 의사결정나무를 전개하는데 필요한 자료의 수집에서 최종 결과분석까지를 4단계로 나누어 진행한다. 1단계와 2단계까지는 현장에서 수집된 자료의 오류와 목표변수로 지정할 자료를 검토하고 엑셀에

서 사용할 용어, 자료의 크기 등을 결정한다. 3단계와 4단계 를 반복해서 자료의 정확성과 결과를 비교분석한다.

1단계	자료 수집 및 마스터테이블 작성
2단계	데이터의 정제 및 통합
3단계	데이터 모형화 및 변환
4단계	의사결정나무 생성 및 결과 분석

그림 5 단계별 데이터마이닝 처리 절차
Fig. 5 Data Mining Procedure by Step

4.1 분석용 자료 수집 및 마스터테이블 작성

제 1단계에서 분석하고자 하는 자료는 2005년부터 2007년 까지 서울의 강북지역에서 발생한 변압기 고장 실적이다. 표 3에서 보는 바와 같이 전체 데이터는 1,070개이고 고장유 형별로 제작불량, 자연열화, 하자 그리고 낙뢰에 의한 고장 으로 구성되어 있다. 이 자료는 전국의 고장유형과는 관리지 역별 특성에 따라 낙뢰와 같은 자연현상의 점유율이 낮고

표 3 고장 유형별 빈도수

Table 3 Frequency of Fault Pattern

고장유형	빈도	퍼센트	유효퍼센트	누적퍼센트
낙뢰	50	4.7	4.7	4.7
불량	475	44.4	44.4	49.1
열화	425	39.7	39.7	88.8
하자	120	11.2	11.2	100.0
합계	1070	100.0	100.0	-

표 4 고장유형 분석용 변수

Table 4 Variable for Analysis of Fault Pattern

자료종류	변수이름(속성)	변수 내용
발생년도	Year(Ordinal)	1: 05년, 2: 06년, 3: 07년
발생 월	Month(Ordinal)	1: 봄(3~5월), 2: 여름(6~8월), 3: 가을(9~11월), 4: 겨울(12~2월)
발생시간	Time(Ordinal)	1: 00:00~06:00(심야), 2: 07:00~12:00(오전), 3: 13:00~18:00(오후), 4: 19:00~24:00(밤)
발생장소	Locat(Nominal)	1: 주택가, 2: 변화가, 3: 상가
결선방식	Connect(Ordinal)	1: 13 2: 21 1: 43
용량	Capa(Ordinal)	1: 10, 2: 15, 3: 20, 4: 30, 5: 50, 6: 75, 7: 100, 8: 150
변압기종류	Type(Nominal)	1: 일반형, 2: 표준형, 3: 저손실형, 4: 아몰포스형
이용율	Use(Ordinal)	1: 0~10, 2: 11~20, 3: 21~30, 4: 31~40, 5: 41~50, 6: 51~60, 7: 61~70, 8: 71~80, 9: 81~90, 10: 91~100
제작사설립년도	Found(Ordinal)	1: 1~5년, 2: 6~10년, 3: 11~20년, 4: 21년 이상
설치경과년도	Install(Ordinal)	1: 17년, 2: 16년, 3: 15년, 4: 14년, 5: 13년, 6: 12년, 7: 11년, 8: 10년, 9: 9년, 10: 8년, 11: 7년, 12: 6년, 13: 5년, 14: 4년, 15: 3년, 16: 2년, 17: 1년
날씨	Weather(binary)	0: 맑음 1: 비
온도	Temp(Interval)	1: -20~-10, 2: -9~0, 3: 1~10, 4: 11~20, 5: 21~30, 6: 31~40
습도	Humid(Interval)	1: 0~10, 2: 11~20, 3: 21~30, 4: 31~40, 5: 41~50, 6: 51~60, 7: 61~70, 8: 71~80, 9: 81~90, 10: 91~100
강우량	Rain(Interval)	1: 1~10, 2: 11~20, 3: 21~30, 4: 31~40, 5: 41~50, 6: 51~60, 7: 61~70, 8: 71~80, 9: 81~90, 10: 91~100
접지측정	E-Test(Binary)	0 : 미시행, 1: 시행
고장유형	Class(Nominal)	열화 : 자연열화, 불량 : 제작불량, 하자, 낙뢰

과부하로 인한 고장은 발생하지 않았으며, 타 지역에서는 거의 없는 하자가 11%를 차지하고 있다. 특히, 제작불량 비 중이 높은 이유는 지역적인 특성상 설비교체 시기가 늦어져 고장이 발생한 것으로 보인다.

마스터테이블의 작성은 실제 고장이 발생한 자료를 엑셀 로 작성하여 데이터마이닝 프로그램에서 동작하도록 하기 위해 자료종류별 고장유형 분석용 변수 변환표를 표 4와 같 이 작성한다. 목표변수를 어느 것으로 하느냐에 따라 전 개되는 나무모형이 달라지므로 작성법을 명확히 이해하는 것이 중요하다. 본 연구에서는 자연열화, 제작불량, 하자 그 리고 낙뢰를 고장유형의 목표변수로 정하고 표 5와 같이 마 스테이블을 완성하였다.

표 5 마스터 테이블 작성 사례

Table 5 Case of Master Table

발생				결선	용	중	이	제	경	날	온	습	강	접	고
년	월	시간	장소	방식	량	류	용율	작사	과년	씨	도	도	우량	지점	장유
1	4	3	1	1	4	1	7	3	10	0	2	5	0	1	열화
1	4	1	1	1	5	1	4	2	9	1	2	6	1	1	낙뢰
1	4	2	2	3	6	2	5	4	5	0	1	4	0	1	하자
1	4	2	1	3	6	1	8	4	7	1	2	5	2	1	불량
1	1	3	2	1	7	1	4	4	12	0	3	6	0	1	불량

4.2 데이터의 정제 및 통합

제 2단계에서는 데이터가 불순자료에 매우 민감하기 때문 에 데이터의 질을 향상시키고 처리과정의 효율성을 향상시 키기 위해서 상관분석을 통해 데이터간 연관성을 분석하고 불필요한 자료를 제거하였다. 예를 들어 날씨가 맑은 날에 강우량이 기재된 경우 오류로 분류하고 정정하였다.

4.3 데이터 모형과 및 변환

제 3단계에서는 SPSS사의 의사결정나무 3.0을 이용하여 나무깊이는 1, 2, 3단계, 전개되는 사례의 최소수는 부모마디 25개, 자식마디는 1개로 하고 가지치기는 표준 예러 Rule을 선택한다. 이익 차트는 Percentiles 10% 값에서 표시되도록 하고 고장유형별로 불량, 열화, 하자 및 낙뢰의 차트를 형성하여 분석을 시도한다. 본 연구에서는 이익차트, 반응차트, 위험도표에 대해 분석한다. 이 자료는 분석용 데이터 70%, 평가용 데이터 30% 비율로 분할한 후 분석용 데이터를 이용하여 모형을 구축하고 그래프를 이용하여 구축된 모형을 평가한다.

4.4 의사결정나무 생성 및 결과 분석

주상변압기 고장유형에 대한 의사결정나무를 CART 알고리즘을 이용하여 생성하였다. 여기서는 주상변압기의 고장유형을 규명하는 변수가 무엇인가를 찾아내는 것이 가장 중요한 사항이다. 특히 고장유형별 상호 연관성이 어떻게 작용하고 있는지 알아본다. 다음은 단계별 분리결과이다.

1) 첫 번째 단계의 의사결정나무 분리 결과

수집된 자료 1,070개를 의사결정나무로 1단계 전개한 결과 고장유형은 총 785대의 자료 중 제작불량에 의한 고장이 45%인 349대이고, 40%가 자연열화에 의한 고장 그리고 구입후 3년 이내에 고장이 발생한 하자 변압기가 11%인 90대, 낙뢰와 같은 자연현상으로 고장이 발생한 것은 4%인 30대이다. 그림 6에서 보는 바와 같이 의사결정나무 모형에서 제시된 첫 번째 분리는 설치경과 연수를 기준으로 전개되었다. 이것은 최초 주상변압기가 설치된 후 8년을 기점으로 고장이 발생하고 있는데 Node 3에서 보면 8년 이하의 주상변압기 고장 유형은 제작불량에 의한 것이 전체의 약 70%로 나타나고 있어 주상변압기 제작사의 품질관리가 중요하다는 점을 알 수 있다. 다시 말해 주상변압기 부품의 불량이나 제작당시 작업자의 부주의에 의한 것으로 판단된다.

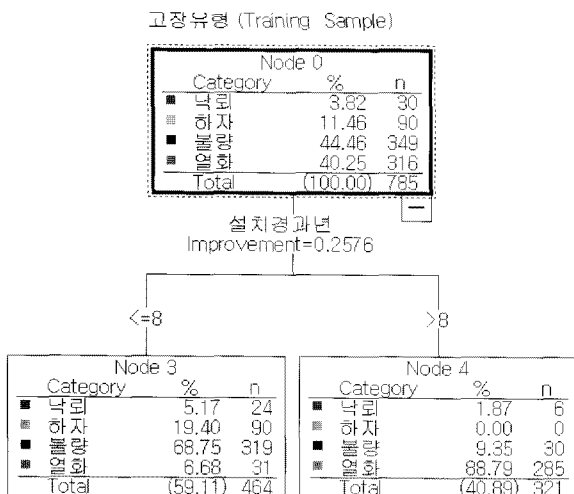


그림 6 첫 번째 나무 분리결과 고장유형
Fig. 6 Fault Pattern of 1st Step Decision Tree

또한 설치 후 3년 이내의 하자처리 대상 변압기도 20%에 달하고 있어 제작시점의 중요성을 보여주고 있다. 뿐만 아니라 낙뢰에 의한 고장도 전체의 5%인 24대로 나타나 사용자재의 내뢰 성능강화가 요망된다.

반면에 열화에 의한 고장은 6.6%를 차지하고 있어 설치 후 8년 동안 운전하는 과정에서 발생한 자재의 불안정성으로 발생한 고장은 상대적으로 적은 것을 알 수 있다.

Node 4에 나타난 결과를 살펴보면 설치 후 8년 초과된 주상변압기의 88.7%가 자연열화에 의한 고장으로 나타났다. 이것은 주상변압기 제작에 사용된 자재와 절연유 등이 일정 시간이 지나면서 급격히 성능이 나빠져 고장이 발생한 것임을 보여주고 있다. 이 수치 속에는 13년 이상된 변압기를 주기적으로 교체하고 있음에도 8년을 기점으로 지속적인 현상순시와 점검으로 열화가 진행되고 있는 변압기를 찾아내어 교체하여야 한다는 것을 의미하고 있다고 보여진다. 제작불량이 전체의 자료의 10% 미만을 차지하고 있는 것은 대부분의 변압기의 자재가 오랜 기간동안 운전해 오면서 안정성을 유지하고 있어 제품의 결함으로 인해 발생하는 고장이 감소한 것으로 보여진다. 이것은 8년이 초과된 주상변압기가 낙뢰로 발생한 고장 점유율이 2% 정도라고 하는 점에서 내구성이 입증된 것으로 판단된다. 주상변압기 설치 초기에 집중적으로 발생하였던 하자 변압기는 3년이 초과된 경우에는 단 1건도 발생되지 않은 것으로 나타나 고장기록 관리가 우수하다고 할 수 있다. 결국 Node 3과 4를 통해 알 수 있는 것은 설치 후 8년 이하와 8년을 초과한 주상변압기의 고장 유형은 제작불량과 자연열화에 의한 고장임을 확인할 수 있었다.

2) 두 번째 단계의 의사결정나무 분리 결과

주상변압기 설치이후 8년을 기준으로 첫 번째 자료 분리 결과를 토대로 나무깊이를 2로 바꾸어 보니 두 번째 단계의 분리결과가 그림 7과 같이 나타났으며 Node의 개수도 증가하였다.

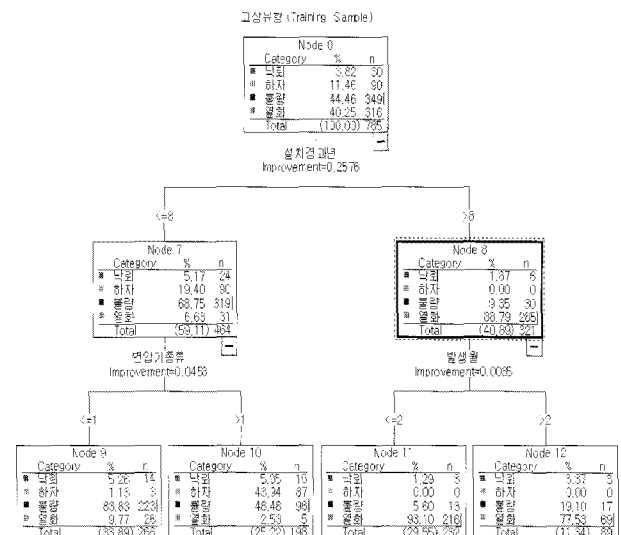


그림 7 두 번째 나무 분리결과 고장유형
Fig. 7 Fault Pattern of 2nd Step Decision Tree

Node 7에서 분리가 일어난 것은 변압기의 종류를 기준으로 하였다. Node 9에서 보면 표 4의 변수에서 보는 바와 같이 변압기 종류가 1보다 작거나 같은 변압기 종류는 일반형 주상변압기이다. 이 Node 9에서는 제작불량이 83.8%를 차지하고 있으며 자연열화와 낙뢰, 하자 순으로 나타났다. 또한 Node 10에서 보면 나머지 표준형, 저손실형, 아몰피스형 등의 주상변압기가 하자와 제작불량으로 각각 43.9%와 48.5%를 나타내고 있다. Node 9와 10을 종합하면 8년 이하의 주상변압기는 대부분 일반형 변압기에서 고장이 발생하고 있다는 것을 알 수 있다. 따라서 제작불량 변압기를 예방하기 위해서는 일반형 변압기의 제작과정에서 철저한 품질관리와 현장에 설치되어 운전 중인 변압기의 정기 점검과 관리 강화가 필요한 것으로 보인다.

Node 8에서 의사결정나무의 분리는 고장발생 월을 기준으로 분리되었다. 고장발생 월은 표 4의 변수에서 발생 월이 2보다 작거나 같은 경우는 봄과 여름인 3월부터 8월에 해당된다. 따라서 Node 11에서 보는 바와 같이 설치된 후 8년 이상 경과되어 고장이 발생한 주상변압기의 93.1%가 자연열화에 의한 것으로 나타났으며, 제작불량과 낙뢰로 인한 고장이 7%를 점유하고 있다. 이것은 주로 습기가 많고 주위 온도가 높은 봄과 여름철에 고장이 많이 발생하고 있다는 것을 보여준다. 다시 말해 설치된 지 8년이 초과된 주상변압기의 자연열화에 의한 고장예방은 봄철 이전에 고장우려대상을 파악하여 사전점검과 진단으로 대처하여야 한다는 의미가 될 수 있다. Node 12는 설치 후 8년이 초과된 주상변압기 중 가을과 겨울인 9월부터 2월 사이에 주로 발생하고 있는 고장 패턴을 보여주고 있다. 즉, 자연열화와 제작불량이 주변온도가 낮은 가을과 겨울에 97%를 나타내고 있으나 전체의 11%인 89대에 불과하고 고장발생 가능성이 봄과 여름에 비해 상대적으로 낮은 점을 고려하면 제품품질이 향상될 경우 모두 해소될 수 있을 것으로 보인다.

3) 세 번째 의사결정나무 분리 결과
 의사결정나무의 깊이를 3으로 하고 다시 나무를 전개한 결과 그림 8과 같이 생성되었다. Node 15와 16에서 분리된 것은 주상변압기의 결선방식과 주상변압기 제작사의 설립경과 연수를 기준으로 하여 나타났다. Node 23에서 보면 표 4의 변수에 따라 결선방식이 1보다 작거나 같은 경우에는 220V를 공급하기 위한 것으로 전체의 8%를 차지하고 있어 비중이 약한 것을 알 수 있다. 한편 Node 24의 경우에는 고장 주상변압기 중 제작불량에 의한 것이 89%로 나타났는데 이것은 380V를 공급하고 있는 변압기에서 발생한 것이다. 결국 일반주택에 전기를 공급하는 주상변압기보다 상가나 소규모 공장과 같은 곳에 공급하고 있는 경우가 부하변동이나 개폐시지로 고장 발생가능성이 더 높다는 것을 의미하고 있다. 또한 Node 16에서 분리된 기준은 제작사가 설립된 후 경과된 연수를 나타내고 있는 것으로 표 4의 변수에서 2를 초과한 경우 제작사가 주상변압기를 생산한 지 10년이 넘었다는 것을 의미한다. 이 Node에서는 하자과 불량률 각각 52%와 47%를 나타내고 있어 오래된 기업에서 생산된 제품이 기술이 축적되어 품질이 우수하다고 판단하기 곤란하다는 결과를 보여주고 있다. 특히, 변압기 종류가 일반형 이외의 제품에서 주로 고장이 발생한 점을 고려한다면 저손실형이나 아몰피스형 변압기와 같은 새로운 제품의 개발 및 설치과정에서 문제를 내포하고 있었던 것으로 추측할 수 있다. 또한 Node 17과 18의 결과에서 고장발생 시간과 고장발생 당일의 습도를 기준으로 분리가 전개되었다. Node 17에서 보면 고장은 부하사용량이 적은 심야보다 사람들의 활동이 집중된 주간과 야간에 주로 발생하고 있으며, 대부분 자연열화에 의한 고장유형으로 나타나 운전경과 연수가 8년이 초과된 변압기는 주로 주간과 야간에 고장이 많이 발생한다는 것을 알 수 있다.

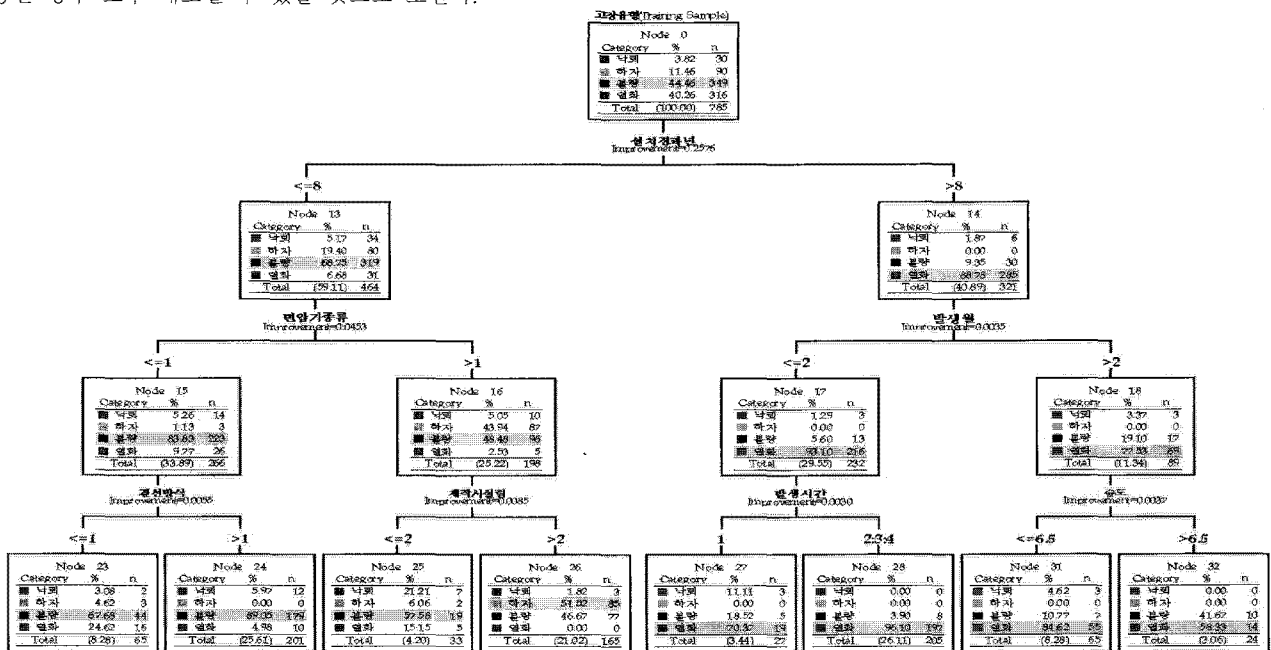


그림 8 세 번째 나무분리 결과 고장유형
 Fig. 8 Fault Pattern of 3rd Step Decision Tree

또한 Node 31과 32에서 보는 것과 같이 고장당일의 습도가 60%를 초과한 경우에 제작불량과 열화가 주로 발생한다는 것도 확인되었다. 세 번째 의사결정나무를 분리한 결과를 종합하면 대부분의 변압기는 8년을 기점으로 설치된 지 8년 이하의 일반형 변압기중 380V의 동력을 사용하고 있는 지역에서 주로 고장이 발생하며, 일반형 변압기 이외의 경우에는 제작사가 설립된 후 10년이 초과된 기업의 제품이 10년 이하의 기업제품보다 6배 이상 고장 발생률이 높은 것을 알 수 있다. 여기에는 그동안 생산된 량의 누적에 의한 것도 작용되었을 것으로 추측된다. 또한 설치된 지 8년이 초과된 주상변압기는 봄철과 여름철에 주로 고장이 많이 발생하며, 심야를 제외한 주간과 야간의 전력사용이 많은 시간에 고장이 발생하고 가을과 겨울철에는 습도가 60%이하인 지역의 변압기가 자연열화로 고장이 많이 발생하고 있음을 알 수 있다. 다시 말해 대부분의 주상변압기는 습도가 많지 않은 계절에 자연열화로 고장이 발생한다는 것을 의미한다. 결국 주상변압기의 고장발생은 여러 가지 유형에 의해 발생하지만 데이터마이닝의 CART 알고리즘에 의해 알게 된 것은 주상변압기의 설치 후 경과 연수와 변압기의 종류 그리고 사용전압에 의한 것과 고장발생의 계절요인과 시간과의 상관관계에 따라 고장이 발생한다는 것을 알 수 있다.

4) 주상변압기 고장유형분류 논리식

그림 8의 의사결정나무를 이용한 주상변압기 고장유형 분리결과를 토대로 고장이 발생하고 있는 주요인의 논리식을 Node 0을 기준으로 작성하면 다음과 같다.

표 6 고장유형 분류 논리식

Table 6 Logic Condition for Classification of Fault Pattern

주요인	Node	조건
열화	28	경과연수>8 \wedge 발생 월 \leq 2 \wedge 발생시간>2
불량	24	경과연수 \leq 8 \wedge 변압기종류 \leq 1 \wedge 결선방식>1
하자	26	경과연수 \leq 8 \wedge 변압기종류>1 \wedge 제작사설립>2
낙뢰	24	경과연수 \leq 8 \wedge 변압기종류 \leq 1 \wedge 제작사설립>2

표 6에서 보는 바와 같이 주상변압기가 고장이 발생하였을 경우 그림 8의 Node 0에서부터 전개되는 의사결정나무를 통해 고장이 발생하였을 경우 설치 경과연수가 8년을 초과하고 발생월은 봄과 여름철 그리고 주로 주간과 야간에 발생하는 고장은 열화일 가능성이 높으며, 경과연수가 8년보다 작거나 같고 변압기 종류가 일반형이면서 공급전압이 220V일 경우에는 불량변압기일 가능성이 높다고 판단할 수 있다. 또한 하자변압기는 경과연수가 8년보다 작거나 같고 변압기 종류가 일반형이 아니면서 제작사가 설립된 지 10년이 초과된 경우에 해당된다고 볼 수 있으며, 낙뢰는 하자과 같은 조건에서 일반형인 경우에 주로 발생하는 것으로 분류할 수 있다.

5) 의사결정나무를 활용한 주상변압기의 고장유형 예측 모델

현행 주상변압기의 고장예측방법은 주로 주상변압기 운전 경과연수가 13년 이상이거나 비가 많이 오고 낙뢰가 치는 경우 고장이 많이 발생한다고 판단해 왔다.

그림 9와 같이 주상변압기 운전 중 고장발생을 예측하기 위해서는 현재 운전 중인 변압기의 자료를 조회하여 13년 이상인 경우를 대상으로 고장이 발생할 확률이 높다고 판정하고, 고장이 발생하지 않은 주상변압기는 별도관리를 통해 시행해 왔다.

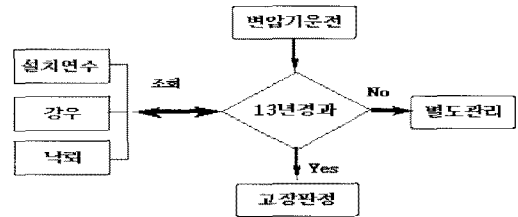


그림 9 고장유형 판정 절차
Fig. 9 Fault Pattern Decision Procedure

그러나 주상변압기가 고장이 발생하였을 경우 의사결정나무를 통해 얻어진 결과를 토대로 고장발생 변압기를 대상으로 그림 10과 같은 고장유형 복구 예측모델을 구축하였다.

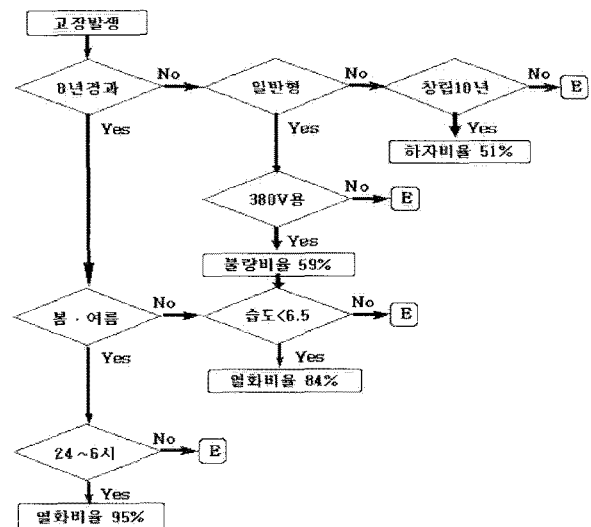


그림 10 변압기 고장유형별 복구 예측모델
Fig. 10 Fault Pattern Restoration Prediction Model for Pole Transformer

다시 말해 변압기 고장 발생은 설치 후 8년이 초과되고 봄과 여름철의 심야 이외의 시간에 발생하는 주상변압기의 고장은 열화에 의한 고장으로 분류가 가능할 것으로 판단되며, 설치 후 8년 이내일지라도 변압기 종류가 일반형이고 380V를 공급하고 있는 변압기의 경우 고장이 발생하면 불량으로 분류할 수 있다고 예측된다. 다만, 예측의 정확도를 높이기 위해서는 다양한 변수의 추가가 필요하다.

6) 고장유형 예측결과를 활용한 고장복구 절차 단축

주상변압기가 고장이 발생하였을 경우 의사결정나무에 의해 고장유형을 판별하게 되면 고장이 발생한 장소나 시간에 따라 고장복구에 필요한 기자재를 미리 가지고 현장에 이동함으로써 고장복구 절차의 단축으로 정전된 사용자의 불편을 감소할 수 있을 것으로 전망된다.

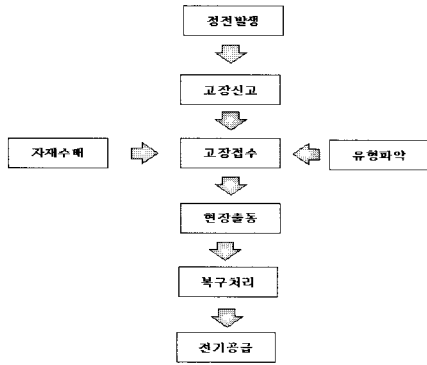


그림 11 개선된 고장복구 절차
Fig. 11 Improved Fault Restoration Procedure

7) 의사결정나무에서 얻어진 각종 그래프

의사결정나무에서 구할 수 있는 도표는 이익차트, 리프트 차트, 반응차트가 있다. 이번 연구를 통해 얻어진 주상변압기 고장유형 중 하자처리와 관련된 이익차트는 아래와 같다. 차트에서 보는 바와 같이 가로축의 고장발생 반응율이 약 20%에 이르렀을 때 세로축의 반응율이 90%를 나타내 모형성능이 우수하다고 할 수 있다.

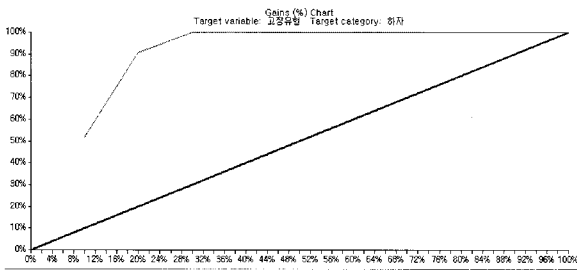


그림 12 이익차트
Fig. 12 Gain Chart

또한 반응차트는 각 구간별 퍼센티지 구간에 대한 'Resp : %' 값을 그래프로 나타낸 것이다. 이 차트에서 구축한 모형에서 주상변압기가 열화일 가능성이 높은 상위 30%는 약 95% 정도의 실제 열화가 발생된 것을 포함하고 있다.

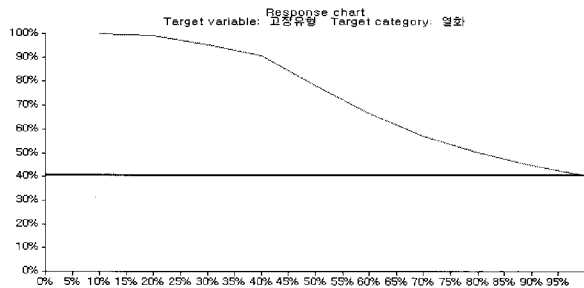


그림 13 반응차트
Fig. 13 Response Chart

이 차트의 기준선은 전체 관찰치중 '열화'가 차지하는 비율이 40%를 나타내고 있음을 알 수 있고 리프트차트와 유사하게 전개된다.

위험도표는 통계학의 관별분석 등에서 알려진 분류행렬로 구축된 모형에 입력 변수 값을 적용하여 사후확률을 얻은 뒤 이 값으로 각 관찰치를 목표변수의 어느 한 범주로 할당함을 나타낸다.

Training Sample						
Misclassification Matrix						
		Actual Category				
		낙뢰	허지	불량	열화	Total
Predicted Category	낙뢰	7	0	0	2	9
	허지	3	52	33	0	88
	불량	17	38	237	34	306
	열화	3	0	19	280	302
Total		30	90	349	316	785

Risk Statistics	
Risk Estimate	0.189809
SE of Risk Estimate	0.0139864

그림 14 위험도표
Fig. 14 Risk Chart

위험도표를 통해 확인할 수 있는 것은 모형의 위험추정치와 이에 대한 표준 오차 값을 알 수 있다. 위험추정치는 (비대각 빈도의 합/ 전체 관찰치의 개수)로 계산되며 위 도표에서는 $(2+3+33+3+19+280)/785 = 0.189809$ 로 계산된다. 이 값은 구축된 모형이 하나의 관찰치를 오분류될 확률이 18.9%라는 것일 나타낸다. 위험추정치의 표준오차는 위험추정치가 이항분포의 성공확률을 나타내는 모수로 적용될 경우 추정치에 계산되는 표준오차 산출과 동일하게 되어 0.017732로 계산된다.

3. 결 론

본 연구에서는 의사결정나무를 이용하여 배전선로에서 전력공급용으로 사용되고 있는 주상변압기 고장유형을 분석하고 유사조건 하에서 주상변압기 고장이 발생하였을 때 고장유형을 예측할 수 있는 모형과 복구절차를 단축하여 사용자의 편의를 제공할 수 있는 방안을 도출하였다. 주상변압기의 고장자료는 실제 현장에서 취득된 자료를 이용하여 정제하고 의사결정나무를 통해 주상변압기의 주요 고장유형을 제작불량, 자연열화, 제작된 제품의 하자 그리고 낙뢰로 분류하였다. 즉 주상변압기는 설치된 후 8년 이내의 경우에는 주로 일반형 변압기가 동력을 사용하는 사용자 지역에서 고장이 발생하고 있다는 것이다. 또한 8년이 초과된 변압기는 봄과 여름의 기온이 높은 시기에 주로 오전, 오후 그리고 야간에 고장이 발생한다고 하는 것이다. 이것은 그동안 주상변압기는 비나 눈과 같이 기상이 나쁘고 주로 심야에 많이 발생한다고 하는 것과는 다른 결과를 보여주고 있다. 이러한 결과를 토대로 향후 주상변압기의 제작 시 내구성이 높은 재질을 사용하고 제품의 조립과정에서 꼼꼼한 마무리로 품질을 높일 경우 제작불량과 같은 고장유형은 크게 감소될 것으로 판단된다. 또한 주상변압기를 운전하고 관리하는 입장에서 기온이 상승하는 봄과 여름철 그리고 주간과 야간에 고장이 많이 발생하는 점에 착안하여 기온이 상승하는 봄철 이전에 주상변압기의 이상 유무를 점검하여 불량유려 변압기를 교체하고, 주로 고장이 많이 발생하는 100kVA미만의 주상변압기를 예비품으로 충분히 확보하여 돌발적인 정전 시 현장에 가지고 이동하여 복구에 활용할 경우 사용

자의 불편시간을 단축할 수 있을 것이다. 결국 주상변압기는 8년 이내에는 주로 동력 사용 장소에서 일반형 변압기가 고장이 발생한다는 것과 운전기간이 8년이 초과된 경우에는 주변 기온이 높아지는 시기와 낮 시간대에 주로 고장이 발생한다는 사실이 고장유형 분석을 통해 입증되었다. 특히 외부 환경요인과 내부 열화요인과의 상관관계가 주상변압기 고장발생에 어떻게 작용하고 있는지를 데이터마이닝기법을 이용하여 규명한 것은 커다란 성과라 할 수 있다. 이번 연구에서는 고장유형 자료를 취득지역으로 국한하여 분석하였기 때문에 타 지역에서 발생하고 있는 과부하와 강풍, 폭설 등과 같은 자연현상으로 주상변압기의 고장이 발생하는 유형은 포함되지 않아 일반적인 결과라고 예측하기는 어렵다. 향후 보다 더 많은 자료를 전국을 대상으로 수집하여 유사 조건하에서 고장이 발생할 경우 주상변압기의 고장유형을 더욱더 정밀하게 예측할 수 있는 연구가 필요할 것으로 기대된다.

참 고 문 헌

- [1] 한국전력, 배전실무 교육교재, 2004 ~ 2007
- [2] 한국전력(중앙교육원), 배전실무반 교육용 교재, 2007
- [3] 한국전력, '배전설비고장분석 및 예방대책', 2004 ~ 2007
- [4] 김경열 외 4인, 분광 광도법 및 전기분석법을 이용한 절연유의 경년열화 진단에 관한 연구, 조명 전기설비 학회 제12권 제2호, pp 15 ~ 20, 1998
- [5] 박희로 외 4인, "가스분석에 의한 주상변압기 예방진단에 관한 연구", 조명전기설비학회지, 제10권 제5호 pp 94 ~ 99, 1996
- [6] 송현진, "주상변압기 열화진단을 위한 원격 측정", 조명전기설비 학회 논문지 제17권 제 1호, 2003
- [7] A.S Farag, "Diagnosing Failed Distribution Transformers Using Neural Networks", IEEE Transactions on Power Delivery, Vol. 16, No. 4. pp 631 ~ pp 636, 2001
- [8] H. Manjari Dola, "Data mining for Distribution System Fault Classification", Power Symposium, Proceedings of the 37th Annual North American, IEEE, pp 457 ~ 462, 2005.
- [9] Young-Han Yoon, "Development of A Deterioration Diagnosis Device For Pole Transformer Using Signal Processing And Wireless communication", IEEE, Power Engineering Society Summer Meeting, pp 1147 ~ pp 1152, 2000
- [10] 최중후 외, Answer Tree 3.0을 이용한 데이터마이닝 예측 및 활용, SPSS 아카데미, 2002
- [11] 배화수 외, "SAS Enterprise Miner를 이용한 데이터 마이닝, 교우사, 2005
- [12] 변성규, "데이터마이닝기법을 이용한 제조공정내의 불량항목별 예측방법, Journal of the Society of KISE, Vol.27 No. 2, pp 10 ~ pp 16, 2004
- [13] 송영석, "의사결정나무모형을 이용한 편마암 지역에서 급경사지 재해예측기법 개발", EG학회지, Vol 18, No. 1, pp 45 ~ pp 54, 2008
- [14] 이형래, " 예측을 향상을 위한 데이터마이닝 기법에 관한 연구", 국방대학교, 2005
- [15] 임영문, " CART 알고리즘을 이용한 전공지탈자 예측 모형에 관한 연구", 한국산업경영시스템학회, 추계학술대회, pp 259 ~ pp 262, 2006

저 자 소 개



황 우 현 (黃 友 炫)

1960년 1월 27일생. 1983년 중앙대학교 전기공학과 졸업. 2000년 한양대학교 전기공학과 졸업(석사), 1997년~2004년 한전 배전자동화 개발 기획 및 운영 1986 ~ 현재 한국전력공사 부장, 2008 대한전기학회 정회원

Tel : 02-3456-3710
 Fax : 02-3456-3799
 E-mail : hblue@kepco.co.kr



김 자 희 (金 子 姬)

1995년 2월 한국과학기술원 전산학과 졸업. 1997년 2월 한국과학기술원 전산학과 졸업 (석사) 2003년 한국과학기술원 산업공학과 졸업(공학박사). 2004년 비엔나 대학 경영정보학과 방문연구원, 2005년 ~ 현재 서울산업대학교 IT정책전문대학

원 정보시스템 공학과 조교수, 2007년 대한전기학회정회원
 Tel : 02-970-6867
 E-mail : jahec@snut.ac.kr



장 완 성 (蔣 完 成)

1951년 10월 3일생. 1983년 중앙대학교 전기공학과 졸업(학사), 1997년 한양대학교 전기공학과 졸업(석사), 2002 서울대학교 최고경영자과정 졸업 2008 현재 한국전력공사 배전계획처장

Tel : 02-3456-4600
 Fax : 02-3456-4699
 E-mail : dongbag@kepco.co.kr



홍 정 식 (洪 淳 植)

1960년 3월 16일생. 1982년 서울대학교 산업공학과 졸업(공학사), 1985년 서울대학교 대학원졸업(석사), 1988년 서울대학교 졸업(공학박사), 1985~1987 서울대학교 조교 1989 ~ 현재 국립서울산업대학교 산업정보시스템공학과 교수

Tel : 02-970-6474
 E-mail : hong@snut.ac.kr



한 득 수 (韓 得 洙)

1972년 5월 14일생. 2000년 명지대학교 전기공학과 졸업(학사). 2002 ~ 현재 한국 전력공사 근무 2008 대한전기학회 정회원

Tel : 02-901-4274
 Fax : 02-901-4279
 E-mail : dshan@kepco.co.kr