마코프 체인과 고장데이터를 이용한 고장건수 예측에 관한 연구

(A Study on The Prediction of Number of Failures using Markov Chain and Fault Data)

이희태* · 김재철*
(Hee Tae-Lee · Jae-Chul Kim)
숭실대학교*

Abstract

It was accomplished that failure analysis not only failure numbers but also power system components every years. and these informations help power system operation considerably. power system equipment were occurred a break down by natural phenomenon and aging but it was not able to predict this failure number. But many papers and technical repots study for each equipment failure rate and reliability evaluation methods. so this paper show a failure number prediction whole power system component usig Markov theory not each component failure probability. the result present a next month system failure number prediction.

1. 서 론

전력기기의 고장은 자연현상과 같은 임의고장과 수명을 다해 고장나는 열화고장등 다양한 원인에 의해 고장이 발생한다.

이러한 기기의 고장은 축적된 데이터와 확률, 통계이론에 근거하여 고장률, 기기수명평가와 같은 분석을 통해 신뢰도를 확보하려 노력하고 있다.

이러한 분석을 통해서 전력계통의 신뢰성을 해석하기 위한 확률적 접근으로 볼 수 있다.

축적된 고장데이터를 이용하여 전력계통의 신뢰도를 평가 및 해석하고 예측하는 것은 현재의 시스템에 대한 신뢰도를 평가하는 것이다.

전력계통 설비의 고장률을 평가는 일반적으로 가속시험과 같은 데이터를 이용하여 수명을 평가가 가능하고, 또는 누적된 과거의 고장과 관련된 데이터를 이용하여 확률 통계적 방법을 통해 나타낼 수 있다. 대표적인 과거 고장 이력 데이터를 이용한 연구결과 중 참고문헌[1]는 컴퓨터 바이러스의 통계를 이용하여 바이러스 발생위협빈도를 평가하고 이를 토대로 서로간의 상이한 바이러스간의 민감도 분석을 평가 하였다. 본 논문에서는 이를 토대로 전력계통중 변전설비의 데이터를 이용하여 적용하였다.

그리고 참고문헌[2]는 이력 고장데이터를 이용하여 해당 설비의 고장률을 예측하고 평가하였고, 참고문헌[3-4]는 와이블 분포와 통계적 패턴을 통합하여 전자기기의 고장데이터를 이용하여 다음 고장을 일으키는 시간을 예측하는 평가 방법을 분석 하였다.

하지만 이러한 결과들은 고장건수를 예측하는 것이 아니며, 이러한 문헌들은 하나의 기기에 대한 고장이력 데이터를 이용하여 해당 기기에 대한 평가를 의미한다.

개별적인 구성설비의 고장건수는 해당 설비에 국한된 평가만을 나타내므로 시스템 전체에 대한 영향은 아니다. 그래서 마코프 이론을 이용하여 과거 월별 시스템에서 발생한 고장통계 자료를 이용하여 향후 시스템 전체에 발생될 고장건수를 평가하였다.

본 논문에서는 전체 변전 시스템에 대한 고장건수를 가지고 모든 설비에서 고장을 가능하게 하는 횟수를 평가하였다.

하나의 대상설비를 위한 위와 같은 논문과 보고서는 많지만 전체 시스템에 대해 익월에 대한 고장건수 예측은 미비한 실정이다. 전력계통은 여러 가지의 설비의 총 집합체이기 때문에 이러한 설비가 고장모드에 의해 나타날 수 있는 영향을 평가하는 것도 중요하다.

2. 고장건수 데이터 분석

본 논문에서 사용한 데이터는 국내 변전소의 고장통계 자료를 이용하였다.

고장통계 데이터 2000년부터 20007년까지의 56개월간의 월별 고장건수를 근거로 하였으며 이와 관련 통계분포는 그림 1과 같다.

이 데이터에서 의미하는 고장은 순간정전과 영구정전을 야기하는 모든 고장에 대한 건수를 의미한다

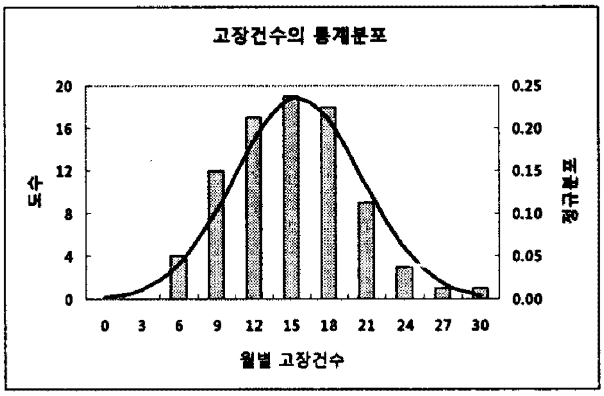

그림 1. 누적된 고장데이터의 통계분포
Fig. 1. Probability distribution of cumulative failure data

그리고 변전설비를 구성하는 설비를 기준으로 하였으며, 고장원인은 제작분량, 시공불량, 보수불량, 작업과실, 열화, 외물접촉, 오동작, 고장파급, 기타등으로 분류하였다.

특히 초기고장에 의한 시공불량과 제작불량을 제외한 고장원인중 그림 2와 같이 기기의 노후에 의한 경년열화와 설비의 오동작에 의한 고장건수가 많은 것으로 조사되었다.

또한 계절에 의한 고장건수는 여름과 가을에 많이 발생한 것으로 조사되었다. 이 기간에 장마와 태풍과 같은 자연현상의 영향으로 생각된다.

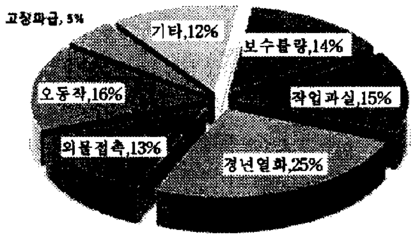

그림 2. 8년간 고장원인별 비율
Fig. 2. Failure cause ratio for 8 years

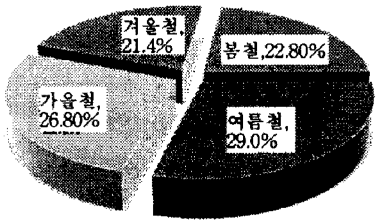

그림 3. 8년간 계절별 고장 비율
Fig. 3. Seasonal failure ratio for 8 years

3. 고장건수 예측 알고리즘

이전의 고장건수 데이터를 통해 향후 시스템의 전체 고장건수를 예측하기 위해 다음 그림 4와 같은 순서를 통해 예측하였다.

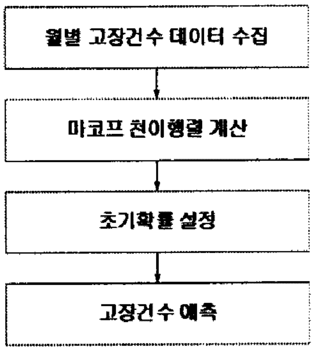

그림 4. 고장건수 예측 순서
Fig. 4. Flow of prediction for the number of failure

3.1 마코프 이론

과거의 고장데이터를 이용하여 향후 발생될 고장건수를 예측하기 위하여 마코프 이론을 이용하였다.

이 이론은 상태(state)간 전이는 이전의 상태에 의존하여 이루어지이며, 향후의 다음 상태를 결정하는 확률 모델이다.

즉 i상태에서 k상태로 전이될 확률은 천이행렬 P와 초기 상태확률 A(0)에 의해 다음과 같이 된다[5-6].

$$P=\begin{bmatrix} p_{11}\ p_{12}\cdots p_{1n} \\ p_{21}\ p_{22}\cdots p_{2n} \\ \vdots\ \ \vdots\ \ddots\ \vdots \\ p_{n1}\ p_{n2}\cdots p_{nn} \end{bmatrix} \tag{1}$$

$$\sum_{j=1}^{n} p_{ik}=1, \quad i=1,2,\cdots,n \tag{2}$$

초기확률 A(0)을 다음과 같다면,

$$A(0)=\{a_1(0),a_2(0),\cdots a_n(0)\} \tag{3}$$

$$\sum_{j=1}^{n} a_j(0)=1 \tag{4}$$

여기서 초기확률 행렬을 선정하기 위해서 고장통계자료를 바탕으로 마지막 1개월부터 3개월까지 나누어 계산하였다.

마지막으로 발생예상건수를 위해서 월별 고장데이터의 값을 초기확률과 같은 방법으로 1개월 데이터뿐만

이 기간에 대한 데이터를 이용한 이유는 수집된 데이터를 어떠한 범위에 따라 결과가 다르기 때문에 범위를 선택하여 평가하였다.

1개월에 대한 데이터만을 고려한 결과는 결국 최대 전달의 고장건수가 최대로 나오기 때문에 사용되는 데이터의 범위를 여러 달로 하여 분석하였다.

4. 고장예측 결과

수집된 고장데이터를 토대로 각각의 상태를 5-state 정의하였고 총 96개월의 데이터의 분포를 계산하고, 계산된 고장건수를 동등한 천이율을 나타내기 위해서 5등분으로 분할하였다.

5-state의 정의는 다음과 같이 정의하였다.

state 1은 월별 고장건수가 9건 미만인 state이고, state 2는 월별 고장건수가 10건 이상 12건 미만인 state, state 3은 12건 이상 14건 미만인 state, state 4는 15건 이상 17건 미만인 state, state 5는 월별 고장건수가 17건 이상인 state이다.

표 1은 년도별 월에 따른 각각의 정의된 state에 대한 값을 표시한 것으로 state의 숫자를 표기 하였다.

표 1. 년도, 월에 따른 각각의 state
Table 1. Each state according to the year and month

년도 / 월별	00년	01년	02년	03년	04년	05년	06년	07년
1월	1	3	4	5	1	1	2	1
2월	1	2	3	2	1	1	3	1
3월	1	1	5	5	3	4	4	1
4월	4	1	3	2	1	1	4	2
5월	4	3	5	3	2	3	2	4
6월	1	3	3	5	3	2	5	4
7월	5	5	5	4	5	3	5	2
8월	5	2	5	5	2	4	5	5
9월	1	1	4	5	5	3	2	2
10월	5	3	5	4	5	3	1	2
11월	2	1	5	3	4	5	3	4

표 1에 state를 토대로 각각의 상태전이를 개수를 E라 하면 식(5)와 같이 나타낼 수 있고, 이를 각각의 확률로 천이행렬 P로 표현하면 식 (6)과 같이 된다.

$$E=\begin{array}{c c}
 & \begin{array}{ccccc} s_1 & s_2 & s_3 & s_4 & s_5 \end{array} \\
\begin{array}{c} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array} & \begin{bmatrix} 5 & 1 & 4 & 2 & 3 \\ 5 & 0 & 4 & 0 & 2 \\ 2 & 2 & 3 & 2 & 7 \\ 3 & 1 & 3 & 1 & 3 \\ 1 & 6 & 2 & 6 & 5 \end{bmatrix}
\end{array} \qquad (6)$$

$$P=\begin{array}{c c}
 & \begin{array}{ccccc} s_1 & s_2 & s_3 & s_4 & s_5 \end{array} \\
\begin{array}{c} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array} & \begin{bmatrix} 0.33 & 0.07 & 0.27 & 0.13 & 0.20 \\ 0.33 & 0.00 & 0.44 & 0.00 & 0.22 \\ 0.13 & 0.13 & 0.19 & 0.13 & 0.44 \\ 0.27 & 0.09 & 0.27 & 0.09 & 0.27 \\ 0.05 & 0.30 & 0.10 & 0.30 & 0.25 \end{bmatrix}
\end{array} \qquad (7)$$

위의 식 (7)에 의한 마코프 체인의 모형은 그림 5와 같이 표현할 수 있다.

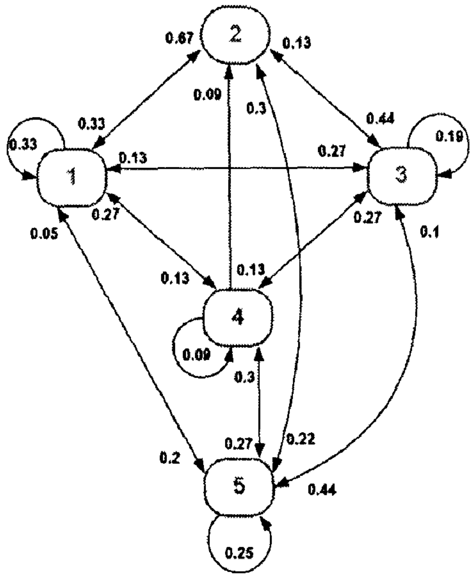

그림 5. 고장건수에 대한 마코프 체인 모형
Fig. 5. Markov chain model for number of failures

다음은 초기행렬 A(0)에 대한 계산과정이다.

초기행렬의 계산을 위해서 데이터의 마지막의 1에서 3개월에 해당하는 state를 알아야 한다.

수집된 데이터의 마지막 1~3월에 해당하는 state는 2, 2, 4 state이며, A(0)에 대한 초기행렬은 식 (8)과 같고, 여기서 윗첨자는 1개월, 2개월, 3개월의 데이터에 대한 행렬을 나타낸다.

$$\{0\ \ 0\ \ 1\ \ 0\ \ 0\ \ 0\},\ A^1(0)=\{\overset{s_1}{0}\ \ \overset{s_2}{0}\ \ \overset{s_3}{1}\ \ \overset{s_4}{0}\ \ \overset{s_5}{0}\ \ \overset{s_6}{0}\} \qquad (8)$$

$$\{0\ \ 0\ \ 1\ \ 1\ \ 0\ \ 0\},\ A^2(0)=\{\overset{s_1}{0}\ \ \overset{s_2}{0}\ \ \overset{s_3}{0.5}\ \ \overset{s_4}{0.5}\ \ \overset{s_5}{0}\ \ \overset{s_6}{0}\}$$

$$\{0\ \ 1\ \ 1\ \ 1\ \ 0\ \ 0\},\ A^3(0)=\{\overset{s_1}{0}\ \ \overset{s_2}{0.33}\ \ \overset{s_3}{0.33}\ \ \overset{s_4}{0.33}\ \ \overset{s_5}{0}\ \ \overset{s_6}{0}\}$$

그리고 천이행렬 P와 초기행렬 A(0)의 곱의 값으로 확률이 계산되며, 1~3개월 실제 고장건수 데이터에 state 확률을 곱해주면 해당되는 향후 고장건수를 확률적으로 예측할 수 있기 되며 $A^1(0)$에 대한 결과는 식 (9)~(10)에 의한 과정을 통해 결

과를 도출할 수 있다.

$$A^1(0) \times P = [0\ 0\ 1\ 0\ 0\ 0] \times \tag{9}$$

$$\begin{bmatrix} 0.33 & 0.07 & 0.27 & 0.13 & 0.20 \\ 0.33 & 0.00 & 0.44 & 0.00 & 0.22 \\ 0.13 & 0.13 & 0.19 & 0.13 & 0.44 \\ 0.27 & 0.09 & 0.27 & 0.09 & 0.27 \\ 0.05 & 0.30 & 0.10 & 0.30 & 0.25 \end{bmatrix}$$

$$= [0.13\ \ 0.13\ \ 0.19\ \ 0.13\ \ 0.44]$$

실제 고장건수에 대한 행렬은 1~3개월에 해당되는 고장건수와 정의된 state에 대한 행렬 R은 식 (10)과 같고 윗첨자는 각 개월에 해당되는것을 의미한다.

$$\begin{array}{cc} & s_1\ \ s_2\ \ s_3\ \ s_4\ \ s_5\ \ s_6 \\ R^1 = & \{0\ \ 0\ \ 14\ \ 0\ \ 0\ \ 0\} \end{array} \tag{10}$$

$$\begin{array}{cc} & s_1\ \ s_2\ \ s_3\ \ s_4\ \ s_5\ \ s_6 \\ R^2 = & \{0\ \ 0\ \ 14\ \ 16\ \ 0\ \ 0\} \end{array}$$

$$\begin{array}{cc} & s_1\ \ s_2\ \ s_3\ \ s_4\ \ s_5\ \ s_6 \\ R^3 = & \{0\ \ 10\ \ 14\ \ 16\ \ 0\ \ 0\} \end{array}$$

식 (9)와 1개월에 대한 행렬일 경우 식 (10)의 값을 이용하여 최종적으로 식 (11)과 같은 예상되는 다음달의 고장건수를 예측할 수 있다.

$$A^1(0) \times P \times R^1 = [0.13\ \ 0.13\ \ 0.19\ \ 0.13\ \ 0.44] \times \begin{bmatrix} 0 \\ 0 \\ 14 \\ 0 \\ 0 \end{bmatrix} \cong 3 \tag{11}$$

이와 같은 방법으로 토대로 1~3개월에 대한 행렬을 계산하여 표 와같은 범위내에 다음 달 고장을 예측하였다.

그 결과 최소 3건에서 최대 6건의 고장이 발생될 수 있을 것으로 분석되었다.

표 2. 과거 고장건수와 초기값에 따른 예측된 고장건수

Table 2. Predicted failure numbers through initial failures and the past number of failures

초기값 \ 건수	건수(전달)	건수(2개월)	건수(3개월)
최근 1개월	3	5	6
최근 2개월	4	5	6
최근 3개월	5	6	6

5. 결 론

수집된 과거 고장데이터와 마코프 이론을 이용하여 2008년도 1월에 발생될 고장건수를 예측하였다.

누적된 고장 데이터를 통한 향후 고장건수 예측은 전력시스템에 대한 접근을 통해 여러 가지 고장원인에 의한 발생건수를 포함하고 있다.

이러한 고장건수 예측은 전력계통 시스템에서 발생될 수 있는 고장을 예측할 수 있는 모델로 활용 가능할 것으로 사료된다.

예측된 월별 고장건수 통계에는 설비의 유지보수 또는 교체 등에 의한 영향을 수치로 나타내기 때문에 각각의 state의 이동은 이러한 정보를 모두 지녔다고 볼 수 있다.

감사의 글

본 연구는 지식경제부의 지원에 의하여 기초전력연구원(R-2007-2-055) 주관으로 수행된 과제임.

참 고 문 헌

[1] 김영갑, 백영교, 인호, 백두권, "마코프 프로세스에 기반한 확률적 피해 파급 모델", 한국정보과학회 논문지, Vol. 33, No 8, pp. 524-535, 2006, 8.

[2] Wenyuan Li, "Evaluating Mean Life of Power System Equipment With Limited End-of-Life Failure Data", IEEE Trans. on Power system, Vol. 19, No. 1, pp.236~242, 2004. 2.

[3] Kevin fizgibbon, Ron Barker, Tige Clayton, Nathan Wilson, "A Failure-Forecast Method Based on Weibull And Statistical Pattern Analysis", IEEE Proceedings Annual Reliability and Maintainability Symposium, pp. 516-521, 2002

[4] W .Li, "Incorporating aging failures in power system reliability evaluation," IEEE Trans. Power Syst., vol. 17, pp.918 - 923, Aug. 2002.

[5] Athanasio Papoulis, s. Unnikrishna Pillai, "Probability, Ramdom Variables, And Stochastic Processes 4th", McGrow Hill, 2002

[6] Wenyuan Li, "Risk Assessment of Power Systems", IEEE press, 2005