

Event Tree하에서 베이지안 기법을 이용한 신뢰도예측*

Bayesian Reliability Prediction under Event Tree[†]

박철순*, 전치혁*, 양희중**, 장수영*
C.S.Park*, C.H.Jun*, H.J.Yang**, S.Y.Chang*

ABSTRACT

When modeling a complex system we use an event tree to analyze propagation of failure. An event tree cannot represent the statistical interrelationships among parameters, but it can be represented as a statistically identical influence diagram so that parameter updating can be easily performed. After updating parameters we can calculate posterior distribution of the failure rate for each path. But exact distribution requires considerably complex numerical integration. We propose an approximation method to calculate the posterior and derive the predictive distribution of the time to next failure. Finally we introduce the system which implements our methodology.

1. 서론

신뢰도 예측문제는 장애에 관측될 랜덤 사상의 불확실성을 취급하는 특수한 통계적 추론 문제로 생각될 수 있다. X 를 다음 번 고장까지의 시간을 나타내는 변수, I 를 입수가능한 모든 정보라 할 경우 $P(X|I)$, 즉 모든 입수가능한 정보에 근거를 둔 X 의 예측분포에 대한 설정은 신뢰도 예측의 핵심이라 할 수 있다. 주로 평균고장률(λ)과 같은 점추정량을 구하는데 중점을 두고 있는 재래식방법에서는 충분한 자료를 입수할 수 없는 경우는 부정확할수 있고 또 점추정치만으로는 대상시스템의 신뢰성을 충분히 결정하기가 어렵다. 따라서 유사한 분야에서 축적된 전문가의 지식을 모델 파라미터분포추정에 삽입할 수 있고 점추정

치만이 아닌 관련된 분포를 구함으로써 추후 고장발생에 대한 충분한 정보를 얻을수 있는 베이지안기법이 보다 적절함을 알 수 있다. 복잡한 대상 시스템을 모델링할때 Event Tree를 이용하게 되면 사고의 진행방향을 논리적이고 효과적으로 분석할 수 있고 통계적으로 동일한 영향도(Influence Diagram)로 표현할 수 있으며 이를 통해서 사후분포로써 파라미터보정을 용이하게 수행할 수 있다. 이런방법으로 각 경로파라미터를 보정한후 각 경로에 대한 고장률의 분포및 다음 고장까지의 시간에 대한 예측분포를 구할 수 있다.

Oliver and Yang(1990)은 가지(branch)가 두개씩(성공/실패)인 다단계

* 포항공과대학 산업공학과

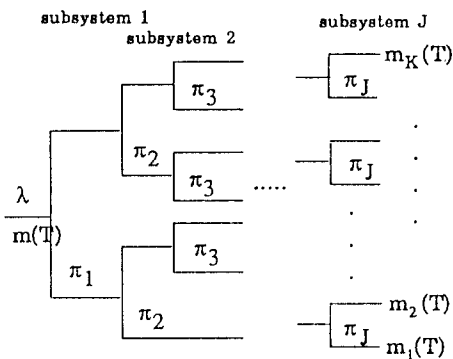
** 청주대학교 산업공학과

+ 본연구는 국방과학연구소 장기기초과제(90-8-1)에 의해 지원되었음.

로 구성된 Event Tree에서 각 가지의 실패 확률을 영향도를 이용하여 베이지안 보정 하는 방법론을 제시하였으며, 전치혁등(1993)은 한단계에서 여러 가지로 구성된 Event Tree에서 각 가지(고장경로)에 해당되는 고장률 및 고장까지시간의 분포를 유도하였다. 본연구에서는 시스템 또는 관심대상인 기능의 고장형태가 Event Tree로 나타날 경우 각 고장경로에 대한 고장률의 사후분포를 근사적으로 산출하고 이를 바탕으로 다음고장까지의 시간에 대한 예측분포를 유도한다. 마지막으로 파라미터 보정작업, 사후분포계산, 고장시간에 대한 예측분포의 계산, 그리고 신뢰도예측으로 이어지는 소프트웨어 시스템을 구성한다.

2. Event Tree

복잡한 시스템에서 사고의 진행방향을 논리적이고 효과적으로 분석하고 Sub-system의 작동여부가 사고의 치명도에 어떻게 영향을 주는가를 파악하기 위해서 Event Tree를 사용한다.



〈그림 1〉 일반적인 Event Tree

〈그림 1〉과 같은 Event Tree를 고려해 볼 때 이 Tree에서 초기이상상태 Event로 시작되어 이 초기이상상태가 조기에 교정되면

시스템에 아무런 영향을 미치지 않던지 경미한 사고로 끝나 버리지만 초기이상상태가 진전되면 후속안전장치의 작동여부에 따라서 발전되어 가는 사고의 경로를 쉽게 볼 수 있다. 위의 그림에서 좌측 첫가지가 초기 이상상태를 나타내고 λ의 율로 발생한다고 본다. Event Tree는 첫가지로 부터 좌측에서 우측으로 진행되어 가며 후속안전장치(subsystem)의 작동여부에 따라서 두가지로 갈라질 수 있으며 해당안전장치에 아무런 영향을 받지 않는 경우는 통과한다. j번째 안전장치에 대하여 윗가지는 정상적으로 가동한 경우, 아랫가지는 작동하지 못하고 실패한 경우를 나타내고 조건부 확률 π_j 가 실패할 확률을 나타내어 아래가지에 할당된다. 〈그림 1〉에서는 후속안전장치가 J개 있는 경우의 Event Tree인데 K가지의 사고경로를 보여주고 있다. Event Tree의 각 경로 끝에 표시된 $m_k(T)$ 는 시간간격 (0, T) 사이에 해당경로를 통해서 발생한 사고수를 나타내고 아래에서 위로 갈수록 치명도가 감소해서 맨윗경로인 $m_k(T)$ 가 가장 경미한 사고수를 나타내고 있다. 일반적으로 Event Tree 모델의 초기이상발생률 λ 와 Branch Parameter π 의 추정을 위해서는 Fault Tree기법을 이용할 수 있다. 그러나 이러한 Event Tree에서는 파라미터간의 통계적 상호연관성을 표시할 수 없고 또 각 경로에 대한 사고자료가 입수되었을 경우 이 자료를 이용하여 파라미터를 보정(update)하는 방법을 찾기가 힘들다. Oliver and Yang (1990)에서는 영향도를 이용해서 각 파라미터와 사고자료사이의 통계적 연관성을 규명하여 각 Branch Parameter를 보정해 나가는 방법을 제시하고 있다.

3. 사후확률분포

초기고장은 λ 가 주어졌을 때 포아송과정으로 발생한다고 가정하며, 이에 따라 λ 의 사전분포(prior)를 Gamma(α, β)로 가정하고 π_j 에

대한 사전분포는 Beta(a, b)로 가정하였다.

3.1 파라미터 보정

Oliver and Yang(1990)에 따르면 λ 와 π 의 사후분포는 다음과 같다.

$$\lambda \mid \text{data} \sim \text{Gamma}(\alpha + m, \beta + T)$$

$$\pi_j \mid \text{data} \sim \text{Beta}(a_j + \sum_{k \in K_j(d)} m_k, b_j + \sum_{k \in K_j(u)} m_k)$$

$j=1, 2, \dots, J$

여기서 $m = \sum m_k$ 이며, $K_j(u)$ 는 j 번째 Subsystem의 윗가지들 를 통과하는 사고 경로의 집합, $K_j(d)$ 는 j 번째의 아래가지들 를 통과하는 사고경로의 집합을 나타낸다.

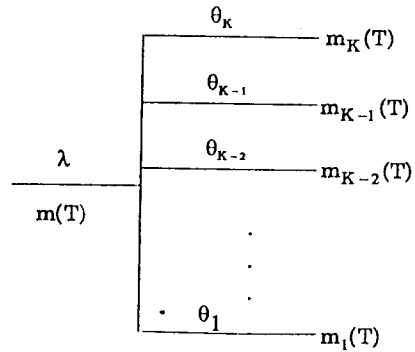
3.2 각 경로에 대한 고장률 분포

초기이상상태발생에 대해 각 경로를 따르게 되는 확률은 각 Subsystem의 작동이 성공했는지 실패했는지의 여부에 따라 π_j 와 $1-\pi_j$, $j=1, 2, \dots, J$, 의 함수로 표시되는데 Event Tree의 k 번째 경로에 해당하는 확률을 θ_k 라 하자. 따라서 $\lambda\theta_k$ 는 Event Tree의 맨 아래로 부터 k 번째 경로를 따라 발생하는 고장률을 나타내며 λ_k 로 표시한다. <그림 1>의 Event Tree는 <그림 2>와 같이 표현될 수 있다. λ_k 의 누적밀도함수 F_k 는 다음과 같이 표시할 수 있다.

$$F_k(z) = P\{\lambda_k \leq z\}$$

$$= \iint_{\lambda\theta_k \leq z} f(\lambda)g(\theta_k)d\lambda d\theta_k \quad (3.1)$$

$f(\lambda)$ 와 $g(\theta_k)$ 는 각각 λ 와 θ_k 의 확률밀도함수이다. 윗식을 z 에 관해서 미분함으로써 λ_k 의 확률밀도함수 $f_k(\lambda_k)$ 이 구해진다. 윗식은 대부분의 경우 closed form으로 표시될수 없으며 <그림 1>과 같이 J 개의 Subsystem이 관련되어있는 경우 $(J+1)$ 차원의 수치적분이



<그림 2> Subsystem을 통합한 예

요구되므로 상당한 컴퓨터시간이 소요되고 또한 수치적분의 근사해를 구하는 과정에서 오차도 피하기 어려우므로 여기서는 감마분포에 근사시키는 방법을 제시한다. θ_k 의 평균을 μ_k , 분산을 σ_k^2 이라 하자. 그러면 λ_k 의 평균과 분산은 아래식과 같이 된다.

$$E[\lambda_k] = (\alpha/\beta)\mu_k$$

$$\text{Var}[\lambda_k] = (\alpha/\beta^2 + \alpha^2/\beta^3)\sigma_k^2 + (\alpha/\beta^3)\mu_k^2$$

σ_k/μ_k 의 값이 작을때는 θ_k 는 근사적으로 상수와 같이 취급될 수 있다. 감마분포를 따르는 확률변수에 상수를 곱해주면 이 역시 감마분포를 따르므로 λ_k 는 근사적으로 파라미터 α_k, β_k 를 갖는 감마분포를 따른다고 볼 수 있다. 즉,

$$\lambda_k \sim \text{Gamma}(\alpha_k, \beta_k)$$

여기서 α_k, β_k 는 다음과 같은 관계를 이용해서 구할 수 있다.

$$\alpha_k/\beta_k = (\alpha/\beta)\mu_k$$

$$\alpha_k/\beta_k^2 = (\alpha/\beta^2 + \alpha^2/\beta^3)\sigma_k^2 + (\alpha/\beta^3)\mu_k^2$$

위의 두식을 정리하면 다음과 같이 α_k, β_k 를 구할 수 있다.

$$\alpha_k = \frac{\alpha \mu_k^2}{(1+\alpha)\sigma_k^2 + \mu_k^2}$$

$$\beta_k = \frac{\beta \mu_k}{(1+\alpha)\sigma_k^2 + \mu_k^2}$$

λ_k 가 위와 같이 감마분포를 따른다고 가정하고 λ_k 가 주어졌을 때 다음고장까지의 시간이 지수분포를 따른다고 가정하면 이 경로를 따라서 발생하는 다음번 고장까지의 시간 X_k 에 대한 예측분포는 closed form으로 간단히 구해질 수 있으며 이는 아래와 같은 Shifted Pareto분포를 따르게 된다.

$$p(X_k)$$

$$= \int p(X_k | \lambda_k) p(\lambda_k) d\lambda_k$$

$$= \int \lambda_k e^{-\lambda_k X_k} \beta_k^{\alpha_k} \lambda_k^{\alpha_k - 1} e^{-\beta_k \lambda_k} / \Gamma(\alpha_k) d\lambda_k$$

$$= \left(\frac{\beta_k}{\beta_k + X_k} \right)^{\alpha_k} \left(\frac{\alpha_k}{\beta_k + X_k} \right)$$

따라서 베이지안 기법을 적용해 예측하는 경우에는 종종 상당히 복잡한 수치적분이 요구되나 위에서 기술한 근사적인 방법으로 예측분포를 아주 용이하게 구할 수 있다. 이 방법은 σ_k/μ_k 이 작을때에 더욱 정확하게 되는데 우리가 자료를 입수하는대로 θ_k 를 보정하게 되므로 이의 분포는 점점 sharp해 저서 σ_k/μ_k 의 값이 작아지고 따라서 근사적인 방법이 더욱 정확성을 띠게 된다. 여기서 μ_k 와 σ_k^2 은 Beta분포를 따르는 확률변수의 곱에 대한 평균 및 분산이므로 용이하게 산출된다.

3.3 여러 경로가 서로 같은 등급의 고장으로 분류되는 경우

<그림 1>의 일반적인 Event Tree에서 각 경로가 서로 다른 유형의 고장으로 분류되는 것이 아니라 그중 여러 경로들은 서로 동일한 유형의 고장으로 취급되는 경우가 종종있다. 예를들어 d, w, h번째 경로를 따라 발생하는 고장은 서로 같은 유형의 고장으로 분류된다고 하자. 그러면 이러한 유형의 고장

은 d, w, h중 어느 경로를 따라도 되므로 이러한 유형의 고장이 발생할때까지의 시간 X는 d, w, h의 경로를 따라 발생하는 고장까지의 시간중 최소치가 된다.

$$X = \min(X_d, X_h, X_w)$$

따라서 이 유형의 고장이 발생할 때까지의 시간 X의 분포는 각 경로의 고장까지시간이 서로 독립인 지수분포를 따른다고 할때 다음과 같이 표시할 수 있다.

$$X | \lambda, \theta_d, \theta_h, \dots, \theta_w \sim \text{Exp}(\lambda_d + \lambda_h + \lambda_w)$$

$\lambda_d + \lambda_h + \lambda_w = \lambda(\theta_d + \theta_h + \theta_w)$ 로 표시할 수 있고 앞절에서 설명된 것과 같이 $\theta_d + \theta_h + \theta_w$ 는 초기 이상상태 발생률인 λ 에 비해 상당히 sharp한 분포를 따르므로 근사적으로 상수로 취급될 수 있다. 따라서 앞절에서 설명된 절차를 따라 $\lambda_d + \lambda_h + \lambda_w$ 의 분포도 또다른 감마분포를 따르는 것으로 간주될 수 있어 어느 특정한 유형의 고장이 발생할때까지의 시간의 예측분포도 마찬가지로 구할수 있다. 만일 $\theta_d, \theta_h, \theta_w$ 의 평균이 각각 μ_d, μ_h, μ_w 이고 분산이 $\sigma_d^2, \sigma_h^2, \sigma_w^2$ 이라 하자. 그러면 이 유형 고장률의 평균과 분산은 다음과 같이 표시할 수 있다.

$$E[\lambda_d + \lambda_h + \lambda_w]$$

$$= (\alpha/\beta) (\mu_d + \mu_h + \mu_w)$$

$$\text{Var}[\lambda_d + \lambda_h + \lambda_w]$$

$$= \left(\frac{\alpha}{\beta^2} + \frac{\alpha^2}{\beta^2} \right) \text{Var}[\theta_d + \theta_h + \theta_w]$$

$$+ \frac{\alpha}{\beta^2} (\mu_d + \mu_h + \mu_w)^2$$

여기서 $\text{Var}[\theta_d + \theta_h + \theta_w]$ 는 각각의 분산과 공분산을 고려하여 산출된다. 따라서 이를 이용하여 감마분포에 대한 새로운 파라미터를 유도할 수가 있고 각 경로별 경우와 마찬가지로

지로 고장률분포 및 예측분포를 구할 수가 있다.

3.4 신뢰도 계산

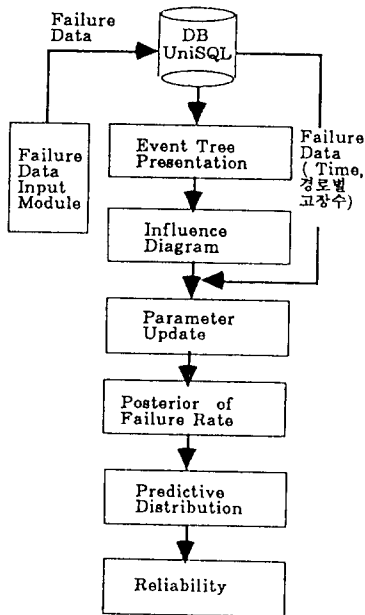
각 경로에 대한 예측분포가 파레토분포를 따르고 있으므로 시간 t 까지 k 에 대한 고장이 발생하지 않을 확률인 신뢰도 $R(t)$ 는 다음과 같다.

$$R(t) = P(X_k > t) = \int_t^{\infty} p(x)dx$$

$$= \left(\frac{\beta_k}{\beta_k + t} \right)^{\alpha_k}$$

4. 시스템 구현 및 실행에

앞에서 언급된 이론을 바탕으로 신뢰도예측 시스템을 구현해 보았다. 데이터베이스는 UniSQL을 이용해서 구현하였고 신뢰도예측시스템의 사용자인터페이스는 X-윈도우환경에서 Motif를 이용해서 구현하였다.

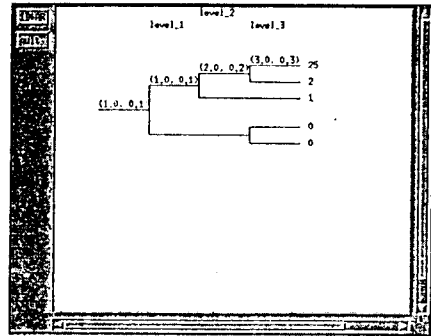


<그림 3> 신뢰도 예측 시스템 흐름도

본 연구에서는 신뢰도예측시스템과 직접적인 관련이 있는 부분에 대해서만 설명하기로 한다. 구현시스템의 전체 흐름도는 <그림 3>과 같고 각 모듈에 대한 설명은 다음과 같다.

4.1 Event Tree

기존의 mission에 대해서 분석을 할 경우는 데이터 베이스에 저장되어 있는 Event Tree를 이용하고 새로운 mission에 대해서 신뢰도를 예측하는 경우는 Event Tree를 새로 만들어서 데이터베이스에 저장후 이용한다. <그림 4>는 선택된 mission에 대한 Event Tree를 표시해 주고 있다.



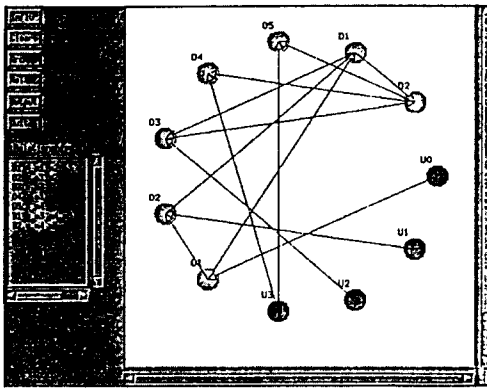
<그림 4> Event Tree 호출

4.2 영향도(Influence Diagram)

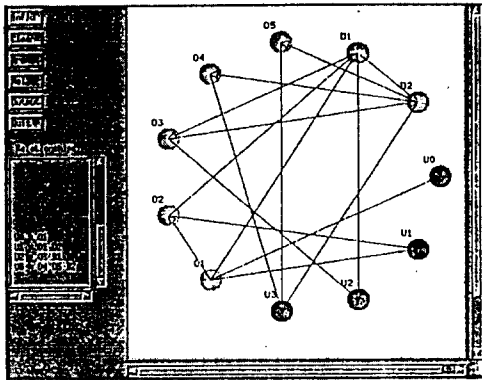
분석할 mission에 대해서 Event Tree를 표시한후 관측불가능한 변수와 관측가능한 변수들의 집합인 각 노드들간의 통계적인 상호연관성을 알아보기 위해서 Event Tree를 통계적으로 동일한 의미의 영향도로 표시해 준다. <그림 5>는 앞의 Event Tree에 대한 영향도를 나타낸다.

4.3 영향도의 Arc Reversal

영향도에서 관측불가능한 변수인 초기이상상태 확률과 각 Subsystem의 작동불량 확률을 나타내는 경로파라미터 π_j 에 영향을 주는 노드들을 찾아내기위해서 즉 관측한 고장데이터를 이용해서 관측불가능한 변수들에 어떻게 영향을 주고 있는가를 알아보기 위해서 Arc Reversal을 수행한다. <그림 6>은 <그림 5>의 Arc Reversal그림을 나타낸다.



< 그림 5 > 영향도

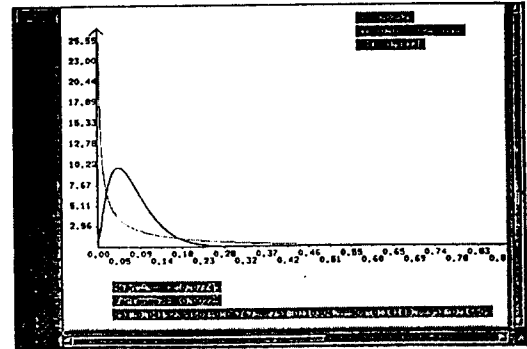


< 그림 6 > Arc Reversal

4.4 파라미터 Updating

고장정보 데이터를 관측한후 초기이상상태 Event를 발생시킬확률 λ 와 각 Subsystem의 작동불량확률인 경로파라미터 π_j 의 파라

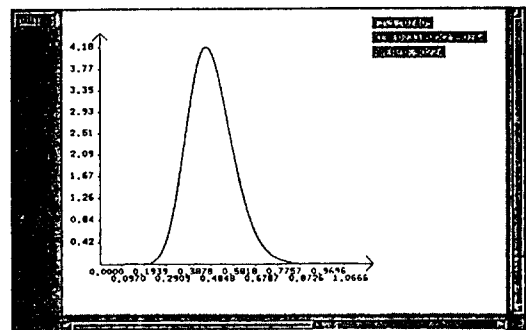
미터를 앞에서와 같은 보정알고리즘을 이용해서 보정해준다. <그림 7>은 한 경로파라미터인 p에 대한 사전분포와 사후분포의 변화를 보여주고 있다.



< 그림 7 > 사전분포와 사후분포비교

4.5 사후확률분포 계산

초기이상상태확률및 경로파라미터를 보정한후에 각 고장 경로별 또는 치명도별 고장률 λ_k 에 대한 파라미터 α_k, β_k 를 근사알고리즘을 이용해서 유도한후 그래프로 표시해서 사용자가 원하는 사후확률값및 여러정보를 알아볼수 있도록 한다. <그림 8>는 사후확률분포를 표시한 예이다.

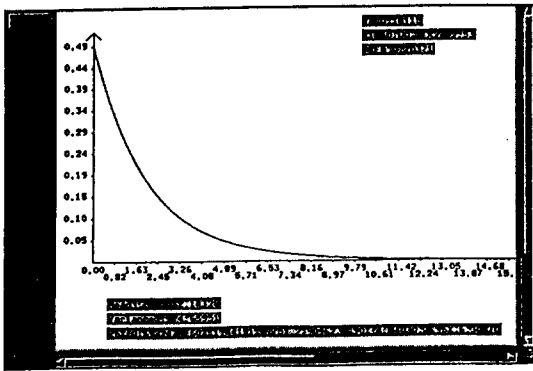


< 그림 8 > 고장률의 사후확률분포

4.6 예측분포계산

각 경로별및 치명도별 구분에 따라서 고장률 λ_k 에 대해 다음 고장까지의 시간에 대한 예측분포를 계산한후 그래프로 표시해 준다.

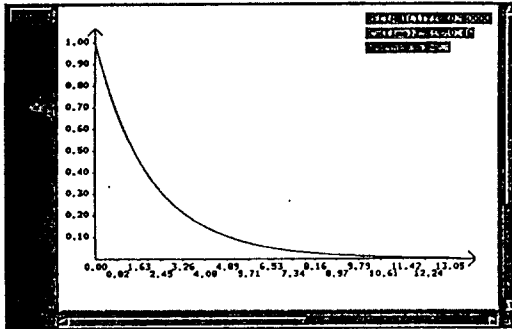
<그림 9>는 앞의 경우에 대한 예측분포를 그래프로 나타낸 예이다.



< 그림 9 > 예측분포

4.7 신뢰도 계산

각 경로별 또는 치명도별 구분에 따라서 고장률 λ_k 에 대해 어느시간까지 시스템이 작동할 확률을 계산해서 <그림 10>과 같이 표시해 준다.



< 그림 10 > 신뢰도

참고문헌

Oliver and Yang (1990), "Bayesian Updating of Event Tree Parameters

to Predict High Risk Incidents" in R.M. Oliver and J.Q. Smith (ed.), *Influence Diagrams, Belief Nets and Decision Analysis*, Chapter 12, Wiley, Chichester, pp. 277-296.

전치혁, 양희중, 장수영, 정의승, 임혜란 (1993), "베이지안기법에 의한 임무신뢰도 예측", *한국경영과학회지*, 18권 1호, pp. 71-78.