

저 빈도 대형 사고의 예측기법에 관한 연구

양 희 중[†]

청주대학교 산업정보시스템공학과

Forecasting low-probability high-risk accidents

Hee-joong Yang[†]

Dept. of Industrial & Information Systems Engineering, Chongju University

We use influence diagrams to describe event trees used in safety analyses of low-probability high-risk incidents. This paper shows how the branch parameters used in the event tree models can be updated by a bayesian method based on the observed counts of certain well-defined subsets of accident sequences. We focus on the analysis of the shared branch parameters, which may frequently often in the real accident initiation and propagation to more severe accident. We also suggest the way to utilize different levels of accident data to forecast low-probability high-risk accidents.

Keywords : Influence Diagram, Bayesian Forecasting, Event Tree

1. 서 론

대형 공장이 안전하게 운행되기 위해서는 여러 하부 시스템들이 우수한 수준의 안전도를 확보 할 수 있도록 제작되고 또 안전 운행 규율에 따라 운행되어야 한다. 이러한 하부 시스템들은 완전 자동화 시스템에 의해 운행될 수도 있고 혹은 인간-기계간의 상호 연관성이 존재하여 작업자의 역할이 시스템 안전성에 크게 영향을 미치는 경우도 있을 수 있다. 이러한 대형 안전 시스템을 설계, 건설하고 운영하는데 PRA(Probabilistic Risk Assessment)가 운용되는 것은 상당히 바람직한 접근법이다[7].

사고가 어떻게 시작되고 어떤 경로를 통해 어떻게 더욱 치명적인 사고로 진행되는지를 체계적으로 설명하기 위해 PRA는 사건 가지(Event Tree)를 종종 이용한다[7]. 원자력발전소 사고와 같은 대형 공장에서의 사고는 일반적으로 그 발생 확률은 작지만 일단 발생하면 피해가 치명적이 되는 특성이 있다[2]. 이러한 사고의 분석에 과거에는 사건 가지 및 결함나무의 분석이 주를 이루고

있었다[3, 5, 9]. 그러나 이러한 분석은 사건 가지를 자료가 입수되는 대로 바로 보정할 수 없다는 단점이 있다. 또한 파라미터간의 상호 독립성 여부도 파악할 수 없다는 단점도 있다. 이에 대한 보완으로 사건 가지를 상호연관도로 변형시켜 모델 파라미터를 보정하는 방법이 연구되었다[10]. 또한 상호연관도 분석에 따른 베이저안 통계기법의 적용 시 초래되는 복잡한 수치적분의 문제를 근사적인 방법으로 단순화 하는 방법도 이미 연구되어 실용적으로 문제해결이 가능해졌다[1, 8].

발생빈도가 낮은 대형 사고(low-probability high-risk accident)는 빈번히 발생하지 않기 때문에 이에 관련된 직접적인 자료 확보가 어렵다. 따라서 피해 규모는 작지만 상대적으로 빈번히 발생하고 있는 다른 유형의 사고 등에 포함되어 있는 정보까지도 최대한 효율적으로 활용할 수 있어야 하며, 또한 일반적인 공학지식과 전문가의견이 효과적으로 반영될 수 있어야 한다. 그러므로 상이한 유형의 사고 자료로부터도 정보를 발췌해낼 수 있는 모델, 전문가 의견 및 과거 운행기록을 포함시킬

[†] 교신저자 hjyang@chongju.ac.kr

수 있는 모델, 자료가 입수되는 대로 즉각 자료 보정이 가능한 모델이 구축되어야 한다. 본 연구에서는 이러한 모델을 구축하고 모데 파라미터를 보정하는 방법을 연구한다. 또한 단순한 사건 가지 분석이 주를 이루었으나 ([6, 7]) 본 연구에서는 기존 연구에서 이루어지지 않았던 공유 파라미터에 관한 분석 방법도 제시된다.

2. 사건 가지(Event Tree)와 결함나무(Fault Tree) 모델

사건 가지나 결함나무와 같은 그래픽 도구들은 발생 가능한 이상 현상들을 열거하고 이들이 더욱 위험한 상태로 진행해가는 과정에서의 상호 연관성을 분석하고, 사고 위험성을 낮추고 전반적인 안전도를 향상시키기 위해서는 어느 시스템이 보완 되어야 하는지 등을 분석하며, 각 이상 상태의 발생 빈도 및 확산 확률 등을 분석하는데 긴요히 활용된다. 또한 하부 시스템의 안전도가 전체 시스템의 안전도에 미치는 영향을 분석하는데에도 활용될 수 있다.

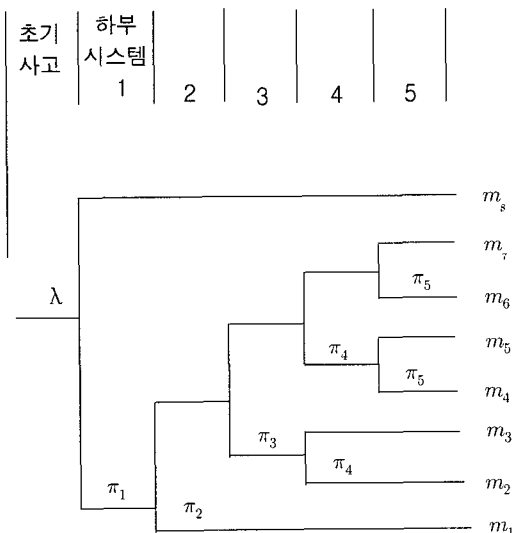
<그림 1>은 대형공장 사고 중 원자력 발전소의 사고를 설명하는 전형적인 사건 가지 중 하나이다. 사건 가지는 위험도가 상대적으로 낮은 초기 이상 상태를 나타내는 가지로부터 시작된다. 실제로 대다수의 초기 이상 상태는 바로 조치되어 공장 운영에 지장을 초래하지 않게 되지만 경우에 따라서는 후속 안전 시스템의 미 작동 혹은 오작동으로 인해 이러한 초기 이상 상태는 더욱 심각한 위험성을 갖는 사고로 진행 될 수 있다. 각 가지는 상단에 표기된 하부 시스템의 성공적 작동 혹은

작동 실패의 경우를 나타낸다. 가지 분기점에서 아래로 갈라지는 경우는 작동 실패를 나타내고 윗 방향으로 갈라지는 경우는 작동 성공을 나타낸다. 아래로 갈라지는 가지에 표기된 π 는 하부 시스템이 실패할 확률을 나타낸다. 이를 가지 파라미터(branch parameter)라고 부른다. $m_1 = m_1(T), m_2 = m_2(T), \dots, m_k = m_k(T)$ 는 각 사고 경로를 따라 (0, T)기간 중 발생한 사고 수를 나타낸다. 일부 하부 시스템은 사건 가지 중의 두 개 이상의 가지에 영향을 주고 있음을 볼 수 있으며 이 경우 우리는 공유 가지(shared branch)가 존재한다고 표현한다. <그림 1>에서 π_4 와 π_5 는 두 가지 경로에 영향을 미치고 있으며 이 경우 π_4 와 π_5 는 가지를 공유하고 있다고 말한다. $m(T)$ 를 다양한 위험도를 가진 벡터형의 사고 수라고 하자. $\delta m(\tau) = m(T+\tau) - m(T)$ 는 (T, T+ τ) 사이에 발생한 사고 증가 개수를 나타내는 벡터, π 를 각 가지의 벡터형 실패확률이라고 하면 우리는 현재 시각 T까지 관측한 자료에 입각해 앞으로 τ 시간 동안 발생할 사고 수를 예측할 필요가 있다. 이는 관측 자료에 입각해 보정된 사후 분포(posterior distribution)와 우도(likelihood)의 곱을 관측 불가능한 파라미터 π 에 대해 적분 소거함으로써 구해진다.

$$p(\delta m(\tau) | m(T)) = \int \dots \int p(\delta m(\tau) | \pi) p(\pi | m(T)) d\pi$$

흔히 신뢰도 분석에서는 모델 파라미터 λ 와 $\pi_j, j = 1, 2, \dots, n$ (n은 하부 시스템의 수)를 관측 자료를 통해 추정하여 특정한 심각성의 사고 발생률 λ_1 를 λ 와 π_j 조합의 곱으로서 구한다. 예를 들어 <그림 1>에서 가장 심각한 사고발생률, λ_1 은 $\lambda \pi_1 \pi_2$ 로 나타내지며 이 값이 관측된 자료를 통해 $\hat{\lambda}_k$ 로 추정되어 특정 유형 사고에 관한 시스템의 신뢰도로 나타내진다. 어느 시스템의 신뢰도 분석은 이러한 사고 발생률의 추정치를 구하는데 집중된다. 그리고 추후 사고가 발생할 때까지의 평균시간은 이 추정치의 역수로서 나타내진다. 그러나 이는 사고까지의 시간이 지수분포를 따르는 특수한 경우에만 적용될 수 있다. 따라서 시스템 신뢰도를 분석하는 또 다른 방법으로서 다음 번 사고가 발생할 때까지의 시간 및 추후 일정 기간 동안의 사고 수에 대한 분포 자체를 구할 수도 있다. 본 논문에서 제시하는 방법은 사고까지의 시간이 지수분포 외에 어떠한 분포를 따르더라도 적용될 수 있다.

가장 위험성이 높은 사고는 보통 상당히 오랜 기간 동안 발생하지 않기 때문에 위험성이 떨어지는 관측 가능한 사고의 수를 이용해 예측해야 한다. 이를 통해 관측되지 않은 유형의 사고에 대한 예측 분포도 시간이



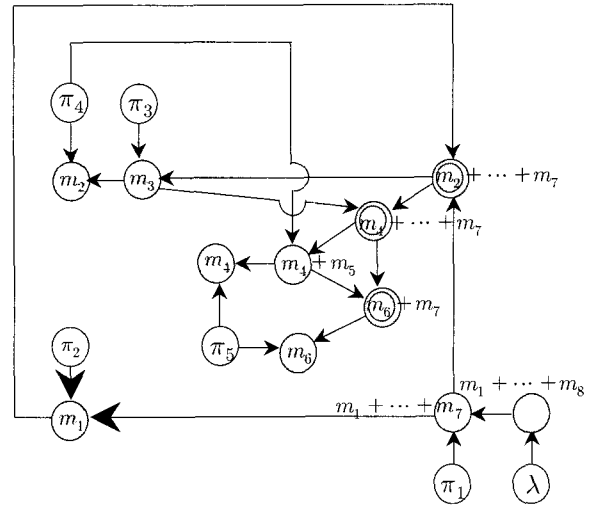
<그림 1> 원자력 발전소사고를 나타내는 사건가지

지남에 따라 그 불확실성을 감소시킬 수 있는 방법이 모색 되어야한다.

사건 가지로부터 통계적으로 동일한 상호연관도를 만들어 가지 파라미터를 보정해가는 방법은 양희중[10]에서 제시되었다. 그러나 여기서는 사건 가지가 후속 시스템의 작동 유무에 따라 한 단계씩 사고 심각성이 높아지는 단순한 사건 가지의 경우에 국한되었었다. 따라서 이러한 사건 가지로부터 만들어진 상호연관도에서는 모델 파라미터가 공유되지 않는 경우만 발생하게 되고 이러한 경우에 대한 파라미터 보정 방법이 연구되었다. 또한 사고를 그 심각성에 따라 여러 형태로 분류하고 각 분류에 따른 사고 예측을 사건 가지를 통해 하는 방법도 연구되었[4]. 그러나 이 경우 역시 사고 심각성이 점진적으로 한 단계씩 확산되는 단순 사건 가지의 경우에 국한되었었다. 본 논문에서는 사고 확산의 형태를 나타내는 사건 가지를 여러 형태의 실제 사고를 반영할 수 있도록 다양화시켜 한 파라미터가 여러 가지에 공유될 수도 있는 경우를 다루고 있다. 이러한 경우에는 모델 파라미터 보정이 더욱 까다로워지지만 공유하고 있는 타 형태의 사고 자료도 이용할 수 있다는 큰 장점도 있을 수 있다. 본 연구에서는 이러한 일반적인 사건 가지를 통계적으로 동일한 사건 가지로 변환시키고 공유 가지를 갖게 되는 모델 파라미터를 효과적으로 보정하고 또한 자료 확보가 어려운 경우 타 유형의 자료를 통해 파라미터를 보정하는 방법을 제시한다.

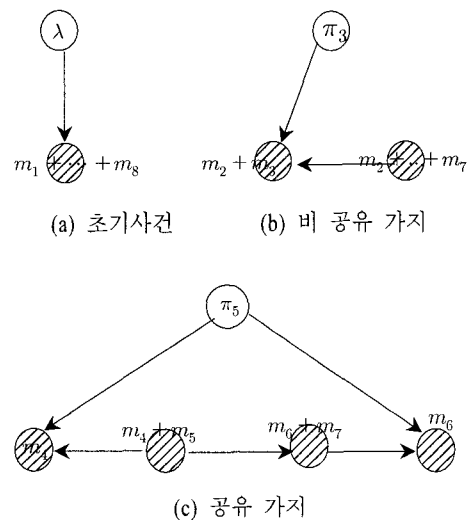
3. 상호연관도 (Influence diagrams)

<그림 1>에 있는 사건 가지는 <그림 2>와 같은 상호연관도로 변형될 수 있다. <그림 2>의 상호연관도는 <그림 1>의 사건 가지와 동일한 파라미터를 공유하고 있다. 초기 이상 상태를 나타내는 node외에 모든 node에는 2개의 선행 node가 있다. 하나는 관측불가능한 가지 파라미터이고 또 하나는 사고수를 나타내는 관측 가능한 node이다. <그림 2>에서 (0, T)기간 동안 발생한 총 사고 수는 $m_1 + m_2 + \dots + m_8$ 이며 이는 초기 이상 상태 발생을 λ 에 영향을 받고 있음을 알 수 있다. 한 하부 시스템의 작동 실패는 고려되고 있는 사고가 사건 가지의 위로 혹은 아래로 진행했는가에 따라 전혀 다른 영향을 미칠 수 있다. 예를 들어 4번째 하부 시스템의 작동 실패는 3번째 하부 시스템의 작동 여부에 따라 전혀 다른 영향을 미치고 있다. 3번째 하부 시스템의 작동 실패는 성공 때와 달리 상당히 심각한 사고로 진행된다.



<그림 2> <그림 1>에 해당하는 상호연관도

상호연관도를 자세히 살펴보면 모델 파라미터들 간에 특별한 구조가 존재함을 알 수 있다. 예를 들어 여러 node중 m_6+m_7 과 π_5 node만이 m_6 에 영향을 미치고 있다. 마찬가지로 m_4+m_5 와 π_5 node만이 m_4 에 영향을 미치고 있음을 볼 수 있다. 이러한 연관 관계는 <그림 3>의 상호연관도에도 나타나 있다. 이 그림은 π 의 직접선행자 (direct successor)와 π 를 보정하는데 필요한 정보를 제공하는 관측 수만으로 구성된 부분 그래프(sub-graph)이다.

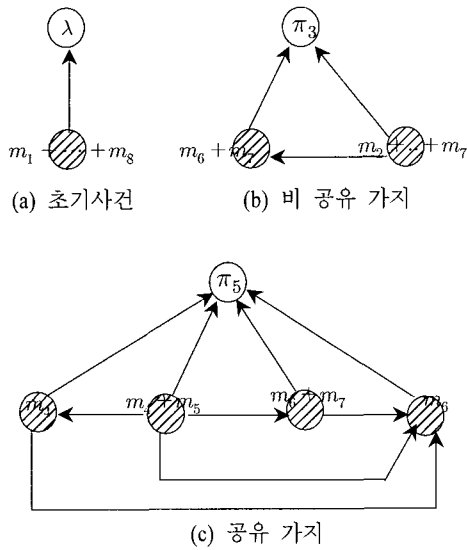


<그림 3> 가지 파라미터와 사건 수에 영향을 받는 node

따라서 <그림 2>의 방대한 상호연관도 모형을 각 파라미터 보정의 측면에서는 소규모의 sub-graph로 나누어 분석할 수 있음을 의미한다.

<그림 3>의 상호연관도를 화살표 방향전환(arc reversal)을 통해 <그림 4>와 같은 상호연관도가 구해지며 이

는 모델 파라미터의 사후 분포를 구하는 과정에 해당된다. 초기 이상 상태 발생율의 사후분포는 총 사고 수에만 영향을 받고 있으며 π_1, π_2, π_3 와 같이 공유되지 않은 가지 파라미터들은 사건 가지의 아래 가치를 지나는 사고 수와 사건 가지의 아래, 위를 지나는 사고 수의 합에 영향을 받고 있다. 한편 π_4, π_5 와 같은 공유 가지 파라미터들은 공유된 가지의 아래 가치들을 지나는 사고 수의 합과 공유된 가지의 아래, 위를 지나는 사고들의 합에 영향을 받고 있다. 이 관계들도 <그림 4>에 도시되어 있다.



<그림 4> 사후분포를 나타내는 상호연관도

4. 사건 가지 파라미터 보정

$\pi_j(j = 1, 2, \dots, J)$ 를 j번째 하부 시스템이 작동 실패할 확률, $m_k(k = 1, 2, \dots, K)$ 를 사건 가지 아래로부터 k번째 경로의 사고 수라 하자. 그리고 d_j^i, u_j^i 를 각각 j번째 하부 시스템 중 아래로부터 i번째($i = 1, 2, \dots, I_j$) 공유 가치를 지나는 사고 수라 하자. 예를 들어 <그림 1>에는 5개의 하부 시스템이 있고 8개의 사고 경로가 있으므로 $J = 5, K = 8$ 이 된다. 그러면

$$\begin{aligned}
 d_1^1 &= m_1 + m_2 + \dots + m_7 & u_1^1 &= m_8 \\
 d_2^1 &= m_1 & u_2^1 &= m_2 + \dots + m_7 \\
 d_3^1 &= m_2 + m_3 & u_3^1 &= m_4 + \dots + m_7 \\
 d_4^1 &= m_2 & u_4^1 &= m_3 \\
 d_4^2 &= m_4 + m_5 & u_4^2 &= m_6 + m_7 \\
 d_5^1 &= m_4 & u_5^1 &= m_5 \\
 d_5^2 &= m_6 & u_5^2 &= m_7
 \end{aligned}$$

가 된다. 데이터의 집합 $D_j(i)$ 를 새롭게 도입하여 이를 아래로부터 j번째 공유 가지상의 작동 실패와 성공 수를 나타내도록 하면 일반적으로

$$D_j(i) = (d_j^1, u_j^1, d_j^2, u_j^2, \dots, d_j^{I_j}, u_j^{I_j}), i = 1, 2, \dots, I_j \text{ 가 되며}$$

$$D_1(1) = (d_1^1, u_1^1) = (m_1 + m_2 + \dots + m_7, m_8)$$

$$D_2(1) = (d_2^1, u_2^1) = (m_1, m_2 + \dots + m_7)$$

$$D_3(1) = (d_3^1, u_3^1) = (m_2 + m_3, m_4 + \dots + m_7)$$

$$D_4(2) = (d_4^1, u_4^1, d_4^2, u_4^2) = (m_2, m_3, m_4 + m_5, m_6 + m_7)$$

$$D_5(2) = (d_5^1, u_5^1, d_5^2, u_5^2) = (m_4, m_5, m_6, m_7)$$

가 된다. 이는 가지 파라미터의 사후 분포를 구하기 위해 사용되는 데이터의 집합을 나타낸다.

여기서 d와 u는 가지의 아래와 위를 지나는 사고 수를 나타내고 있다. 예를 들어 $D_4(2)$ 는 $(m_2, m_3, m_4, m_5, m_6, m_7)$ 와 다르다는 것에 주목해야 한다. 따라서 $D_j = D_j(i)$ 는 π_j 를 보정하기에 충분한 데이터 집합이고 $p(\pi_j | D) = p(\pi_j | D_j(i))$ 가 된다. 그리고 $D = \{m_1, m_2, \dots, m_k\}$ 를 모든 관측 가능한 데이터 집합이라 하고 $m = \sum_k m_k$ 를 초기 이상 상태의 총 발생 건수라고 하면 사건 가지 파라미터의 사후분포는 사전분포가 서로 독립이라는 가정 하에 아래와 같이 각각 파라미터별로 분류되어 구해진 사후 분포들의 곱으로 표시될 수 있다.

$$\begin{aligned}
 p(\lambda, \pi | D) &= p(\lambda, \pi_1, \pi_2, \dots, \pi_J | m_1, m_2, \dots, m_k) \\
 &= p(\lambda | m)p(\pi_1 | D_1(1))p(\pi_2 | D_2(1_2)) \dots p(\pi_J | D_J(I_J))
 \end{aligned}$$

윗 식에서 등호 우측은 아래와 같이 데이터가 관측되는대로 반복적으로 보정된다.

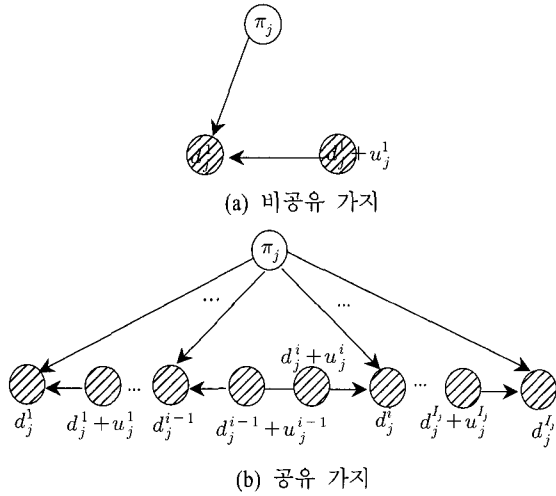
$$\begin{aligned}
 p(\pi_j | D_j(i)) &= C_{ij} p(\pi_j | D_j(i-1))p(d_j^i | \pi_j, d_j^i + u_j^i) \quad i = 1, 2, \dots, I_j
 \end{aligned}$$

$$\begin{aligned}
 p(\pi_j | D_j(1)) &= C_{1j} p(\pi_j)p(d_j^1 | \pi_j, d_j^1 + u_j^1) \quad j = 1, 2, \dots, J
 \end{aligned}$$

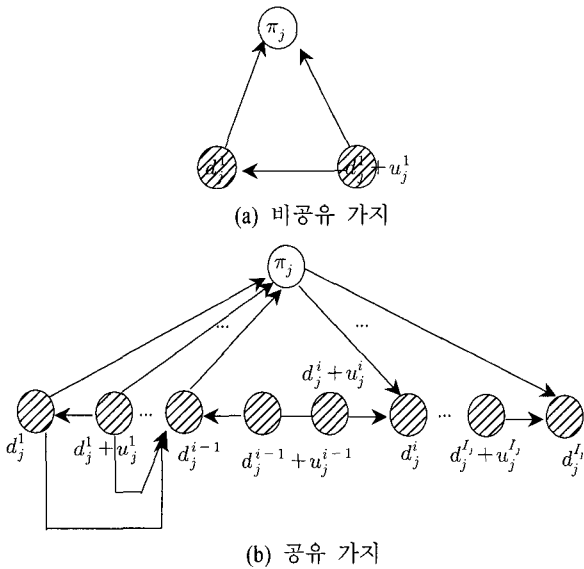
위에 대한 증명은 <그림 5>의 부분 그래프를 보면서 이루어 질 수 있다. <그림 5>(a)는 가지가 공유되지 않은 경우를, <그림 5>(b)는 가지가 공유된 경우를 보여주고 있다.

<그림 5>(a)의 경우에는 π_j 에는 선행 node(predecessor node)가 없고 d_j 가 아닌 $d_j + u_j$ 만이 선행 node를 가질 수 있다. 따라서 π_j 의 사후분포는 우도와 사전분포의 곱, $p(d_j | \pi_j, d_j + u_j)p(\pi_j)$ 에 비례하며 그 외에 다른 모든 종류의 사고 수 혹은 다른 모든 가지 파라미터에 의존하지 않는다. <그림 5>(a)와 같은 비 공유 파라미터의 화살표

방향전환(arc reversal)은 <그림 6>(a)에 나타나 있다.



<그림 5> 가지 파라미터에 영향을 받는 사고 수



<그림 6> 사고 수를 이용한 π 의 보정

이제 <그림 6> (b)와 같이 2개 이상의 가지가 하나의 파라미터를 공유하고 있는 경우를 보자. 이 경우 사후분포는 π_j 의 사전분포와 $\pi_j, d_j^i + u_j^i, i = 1, 2, \dots, I_j$ 가 주어졌을 경우 d_j^i 의 우도에 의존하게 된다. <그림 6> (b)에서 공유 가지에 대한 π_j 의 사후분포는 일반적으로 상당히 복잡한 함수가 될 수 있음을 알 수 있다. 그러나 다행히도 파라미터 보정은 반복적 과정을 통해 간단히 이루어 질 수 있다. 우선,

$$p(\pi_j | d_j^1, u_j^1) = C_1 p(\pi_j) p(d_j^1 | \pi_j, d_j^1 + u_j^1)$$

가 되며 추가적인 데이터 d_j^2, u_j^2 를 이용하여

$$\begin{aligned} p(\pi_j | d_j^1, u_j^1, d_j^2, u_j^2) &= C p(\pi_j) p(d_j^1 | \pi_j, d_j^1 + u_j^1) p(d_j^2 | \pi_j, d_j^2 + u_j^2) \\ &= C_{2j} p(\pi_j | D_j(1)) p(d_j^2 | \pi_j, d_j^2 + u_j^2) \end{aligned}$$

가 된다.

이와 같은 방법으로 공유 시스템에 대한 사고 자료 $D_j(i) = (d_j^1, u_j^1, \dots, d_j^i, u_j^i)$ 를 반복적으로 이용하여 공유 가지 파라미터를 보정할 수 있다.

5. 이항분포시 보정방법

(0, T)기간 중 총 m 개의 초기 이상 상태가 발생하였고 각 가지 파라미터들은 서로 독립이라고 가정되면 8개의 각기 다른 경로를 따르는 사고의 수는 다항분포를 따른다.

$$\begin{aligned} (m_1, m_2, m_8 | m, \pi_1, \pi_5) \\ \sim MN(m, \pi_1 \pi_2, \pi_1(1 - \pi_2) \pi_3 \pi_4, \dots, (1 - \pi_1)) \end{aligned}$$

따라서 어느 특정한 경로의 사고 수는 이항분포를 따른다. 예를 들어 총 사고 수 m 과 가지 파라미터 $\pi_1, \pi_2, \dots, \pi_5$ 가 주어졌을 경우, 사건 가지 아래에서 두 번째 경로의 사고 수 m_2 는 아래의 이항분포를 따른다;

$$(m_2 | m, \pi_1, \pi_2, \dots, \pi_5) \sim \text{Bin}(m, \pi_1(1 - \pi_2) \pi_3 \pi_4)$$

π_j 와 $d_j^i + u_j^i$ 가 주어졌을 때 d_j^i 의 우도가 이항분포를 따를 때는 파라미터 보정이 아주 간단해진다. $K_j(d)$ 와 $K_j(u)$ 를 j 번째 시스템의 아래와 위를 각각 지나는 사고 경로의 집합이라고 하고 $K_j = K_j(d) \cup K_j(u)$ 를 j 번째 시스템의 작동 여부에 영향을 받는 총 사고 경로의 집합이라고 하자.

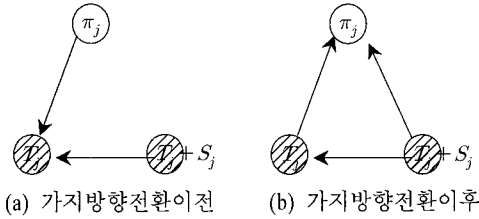
그러면 j 번째 시스템의 아래 가지들을 지나는 사건 수의 합 T_j 와 j 번째 시스템의 윗가지를 지나는 사건 수의 합 S_j 와 아래와 같은 연관성이 성립한다.

$$\begin{aligned} T_j &= \sum_{i=1}^{I_j} d_j^i = \sum_{k \in K_j(d)} m_k, \\ S_j &= \sum_{i=1}^{I_j} U_j^i = \sum_{k \in K_j(u)} m_k \end{aligned}$$

그러면 T_j 와 S_j 는 π_j 를 보정하기에 충분한 정보를 갖고 있기 때문에 이 용어를 이용하면 <그림 5>, <그림 6>에 있는 상호연관도가 <그림 7>처럼 더욱 간결해진다. 이는 이항분포의 특수성 때문으로 간단히 증명될 수 있다. 만약 π_j 가 여러 가지에 공통적으로 영향을 미치고 있다면 데이터 입수후의 사후분포는 j 번째 시스템의 아래 가지들을 지나는 사건 수의 합 T_j 와 j 번째 시스템의

윗가지를 지나는 사건 수의 합 S_j 에만 영향을 받기 때문이다.

$$p(\pi_j|D) = p(\pi_j|D_j(I_j)) = C p(\pi_j) \pi_j^{I_j} (1-\pi_j)^{S_j}$$



<그림 7> 이항분포 우도의 상호연관도

6. 베타 사전분포

π_j 의 사전분포가 파라미터 α_j, β_j 를 갖는 베타분포이고 π_j 와 $d_j^i+u_j^i$ 가 주어졌을 경우 d_j^i 의 우도가 이항분포를 따른다면 π_j 의 사후분포는 사건 가지의 윗 방향과 아래 방향을 지나는 사고 수를 반영하는 새로운 베타분포를 따르게 된다.

$$(\pi_j | D) \sim \text{Be}(\alpha_j + \sum_{k \in K_j(d)} m_k, \beta_j + \sum_{k \in K_j(U)} m_k) \dots \dots \dots (1)$$

여기서 π_j 의 확률 밀도함수 $p(\pi_j)$ 와 기댓값 $E[\pi_j]$ 는 아래와 같다.

$$p(\pi_j) = \frac{\Gamma(\alpha_j + \beta_j)}{\Gamma(\alpha_j)\Gamma(\beta_j)} \pi_j^{\alpha_j-1} (1-\pi_j)^{\beta_j-1},$$

$$E[\pi_j] = \frac{\alpha_j}{\alpha_j + \beta_j}$$

이는 $i = 1, 2, \dots, I_j$ 의 경우에 대해 π_j 와 $d_j^i+u_j^i$ 주어졌을 때 d_j^i 가 이항분포를 따른다면 j 번째 하부 시스템의 공유 가지 중 아래 가지를 지나는 사고수의 총 합은 j 번째 공유 가지를 지나는 총 사고수와 가지 파라미터 π_j 가 주어진 경우 역시 이항분포를 따르게 된다는 사실로부터 입증될 수 있다.

$$\sum_{k \in K_j(d)} m_k \mid \sum_{k \in K_j} m_k, \pi_j \sim \text{Bin}(\sum_{k \in K_j} m_k, \pi_j) \dots \dots \dots (2)$$

따라서 베타 사전 분포의 경우 베타분포의 파라미터 α 항은 공유 가지 파라미터 π 에 영향을 받는 모든 가지 중 아래 가지를 지나는 사고 수의 합이 되고 β 항은 공유 가지 파라미터 π 에 영향을 받는 모든 가지 중 윗

가지를 지나는 사고 수의 합이 된다.

7. 수치 예

<그림 1>의 사건 가치를 따라 사고가 <표 1>에서와 같이 발생했다고 가정하자. 본 논문에서는 파라미터 실시간 보정의 방법을 보여주기 위해 가정된 데이터를 이용한다.

<표 1> <그림 1>의 경로를 따르는 월별 사고수

시간(월)	m_8	m_7	m_6	m_5	m_4	m_3	m_2	m_1
1	90	1	0	0	0	0	0	0
2	200	1	0	0	0	0	0	0
3	310	1	1	0	0	0	0	0
4	440	2	1	1	0	0	0	0
5	480	2	1	1	0	0	0	0
6	630	3	1	1	1	0	0	0
7	750	4	2	1	1	0	0	0
8	820	4	2	1	1	1	0	0
9	910	4	2	2	1	1	0	0
10	1030	5	2	2	2	1	0	0
11	1090	5	3	2	2	1	0	0
12	1250	5	3	3	2	1	0	0

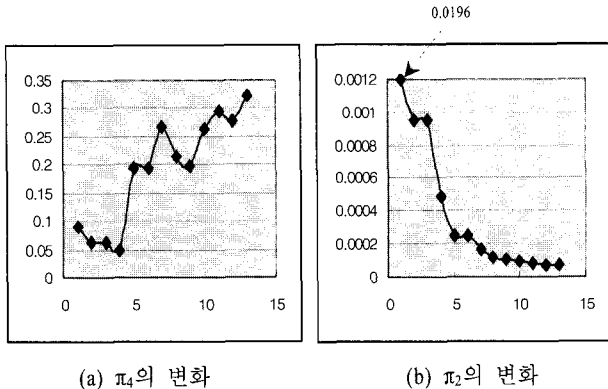
<그림 1>의 5개 sub-system(하부 시스템)의 사고 발생 확률은 베타분포를 따른다고 가정하고 그 사전 파라미터 값들은 <표 2>와 같이 가정된다. 베타분포는 파라미터의 값에 따라 0과 1사이의 거의 모든 모양을 나타낼 수 있는 신축성 있는 분포이기 때문에 확률을 나타내는 사전분포로서 많이 이용된다. 베타분포의 사전 파라미터 값은 과거 자료, 전문가 의견, 시스템에 대한 사전 지식 등을 동원하여 가정되는데 초기 불확실성도 고려하여 적절한 파라미터 값이 가정된다.

<표 2> 가지 확률의 사전 분포 파라미터

	하부 시스템 1	2	3	4	5
α_j	0.001	0.001	0.1	0.2	3.0
β_j	0.1	0.05	1.0	2.0	9.0

이 이후에는 자료가 입수 되는대로 실시간으로 보정된다. 특히 하부 시스템 2와 하부 시스템 4에 대해서 살펴보면 하부 시스템 2는 한 번도 사고가 발생한 적이 없다. 그러나 하부 시스템 4는 가지를 공유하고 있고 이

중 한쪽 경로를 통해서 사고가 없지만 다른 한 쪽 경로를 통해서 사고가 발생하고 있다. π_2 와 π_4 의 평균값을 <그림 6>과 식 (1), 식 (2)에 따라 보정한 결과가 <그림 8>에 나와 있다.



(a) π_4 의 변화 (b) π_2 의 변화
 <그림 8> 시간에 따른 사고 확률 평균값의 변화

하부 시스템 2는 한 번도 사고가 발생하지 않았기 때문에 가지 확률 π_2 의 사후 평균값은 시간이 갈수록 작아진다. 한편 하부 시스템 4의 π_4 에 영향을 받는 사고 m_2 는 한 건도 발생하지 않았으나 이 π_4 는 공유되어 있고 그 공유 가치를 지나는 사고 m_4 와 m_5 는 발생하고 있다. 따라서 π_4 는 π_2 처럼 시간이 감에 따라 지속적으로 감소하는 것이 아니라 등락이 이루어지고 있다.

8. 결론

저 빈도 대형 사고를 나타내는 사건 가치를 상호연관도로 변형 시키는 방법에 대해서 기술하였다. 사건 가치에 포함되어 있는 모든 정보는 통계적으로 동일한 상호연관도로 옮겨질 수 있으며 이에 베이지안 예측 기법을 적용하여 사전분포와 우도를 가정하고 관측된 자료를 통해 모델 파라미터를 적분 소거함으로써 사후 분포를 획득하는 방법을 설명한다. 수학적인 방법과 상호연관도의 그래픽 방법의 동일성 또한 설명되었다. 특히 공유 파라미터를 갖는 사건 가치를 상호연관도로 변형하여 분석하는 방법을 설명하였으며 이 경우 특정 경로로의 사고 발생은 없으나 타 경로를 통해 사고가 발생하더라도 관측 자료로부터 필요한 정보만을 획득하여

모델 파라미터를 보정하는 방법을 설명하였다.

이러한 방법은 어떠한 형태의 사건 가치에도 모두 적용될 수 있기 때문에 다양한 형태로 발생하는 저 빈도 대형 사고의 예측이 훨씬 효율적으로 이루어 질 수 있을 것으로 기대된다.

참고문헌

- [1] Aitchison, I. R. and Dunsmore; "Statistical Prediction Analysis," Cambridge University Press, 1975.
- [2] Bruske, S. Z. and Holland, D. F.; "Risk Assessment Technology for the Evaluation of Tritium Accident Mitigation," *Nuclear Technology Fusion*, 4 : 539-543, 1983.
- [3] Holdren, J. P. et al.; "Exploiting the Competitive Potential of Magnetic Fusion Energy : The Interaction of Economics with Safet and Environmental Characteristics," *Fusion Technology*, 13 : 7-56, 1988.
- [4] Jun, Chi-Hyuck, Chang, S. Y., Hong, Y., and Yang, H. J.; "A Bayesian approach to prediction of system failure rates by cirticalities under event trees," *International journal of production economics*, 623-628, 1999.
- [5] Khan, Faisal et al.; "A New Methodology for Safety Management Based on Feedback from Credible Accident-Probabilistic Fault Tree Analysis System," *Journal of Hazardous Materials*, 87(1) : 23-36, 2002.
- [6] Nuclear, U. S. Regulatory Commission; "Reactor Risk Reference Document," Office of Nuclear Regulatory Research, NUREG-1150, 1-3 : 1987.
- [7] Piet, S. J.; "Implication of Probabilistic Risk Assessment for Fusion Decision Making," *Fusion Technology*, 10 : 31-48, 1986.
- [8] Rasmussen, N. C.; "Method of Hazard Analysis and Nuclear Safety Engineering," *Annals New York Academy of Science* : 20-36, 1981.
- [9] Yang, Heejoong, "Approximation Method in Bayesian Prediction of Nuclear Power Plant Accidents," *한국산업공학회지*, 16(21) : 135-147, 1990.
- [10] Yang, Heejoong; "Forecasting Accidents by Transforming Event Trees into Influence Diagrams," *산업경영시스템학회지*, 29(1) : 72-75, 2006.