

# 마코프 체인을 활용한 산업재해 예측에 관한 연구

박영진<sup>1</sup>, 김태성<sup>2\*</sup>

<sup>1</sup>금오공과대학교 산업공학부 산업경영공학전공 학생, <sup>2</sup>금오공과대학교 산업공학부 교수

## A Study on Industrial Accident Prediction Applying Markov Chain

Yeong Jin Park<sup>1</sup>, Tae Sung Kim<sup>2\*</sup>

<sup>1</sup>Student, School of Industrial Engineering, Kumoh National Institute of Technology

<sup>2</sup>Professor, School of Industrial Engineering, Kumoh National Institute of Technology

**요약** 산업재해중 유형별 가장 많은 재해 건수를 보이는 업종은 건설업과 제조업이다. 이를 예방하기 위한 노력은 꾸준히 지속되고 있다. 본 논문에서는 마코프 체인 프로세스를 도입하여 건설업과 제조업에서의 산업재해 발생을 예측하기 위한 모델링을 제시한다. 이에 사용된 분석자료는 한국산업안전연구원에서 제시하는 산업재해 통계자료로써 월별 자료를 사용하였다. 건설업, 제조업에서의 산업재해에 대한 빈도를 관심(S1), 주의(S2), 경계(S3)로 분류한 뒤 초기확률을 정의한 후, 전이행렬을 곱하여 다음달 예측 가능한 산업재해 빈도를 산출하였다. 본 연구에서는 2014년~2016년까지의 자료를 활용하여 전체에 대한 평균값, 1년 평균값, 최대값으로 분류하여 어느것이 예측률이 더 높은지 2017년 1월 관찰된 실제 데이터와 비교하여 확인하였다. 건설업의 경우 최대값을 적용하였을 때 124의 차이로 가장 작은 차이값을 보여주었고, 제조업의 경우 최대값을 적용하였을 때 108의 차이로 가장 작은 차이값을 보여주었다.

**키워드** : 산업재해, 건설업 산업재해, 제조업 산업재해, 마코프 체인, 예측 모형 모델링

**Abstract** Among industrial accidents, the industries with the highest number of accidents by type are the construction industry and the manufacturing industry. Efforts to prevent this are continuously ongoing. In this paper, we propose a modeling for predicting industrial accidents in the construction and manufacturing industries by introducing the Markov chain process. The analysis data used in this study were statistical data on occupational accidents presented by the Korea Occupational Safety and Health Institute, and monthly data were analyzed. After classifying the frequency of occupational accidents in the construction and manufacturing industries into interest (S1), caution (S2), and boundary (S3), the initial probability was defined, and then the predictable frequency of occupational accidents for the next month was calculated by multiplying it by a transition matrix. In this study, data from 2014 to 2016 were used to classify the overall average value, the one-year average value, and the maximum value to determine which has the higher prediction rate by comparing it with the actual data observed in January 2017. In the case of the construction industry, the smallest difference was shown with a difference of 124 when the maximum value was applied, and in the case of the manufacturing industry, the smallest difference value was shown with a difference of 108 when the maximum value was applied.

**Key Words** : Industrial Accidents, Construction Accidents, Manufacturing Accidents, Markov Chain Process, Predictive Modeling

\*Corresponding Author : Tae Sung Kim (tkim@kumoh.ac.kr)

Received November 02, 2022

Revised December 16, 2022

Accepted December 24, 2022

Published December 28, 2022

## 1. 서론

2017년 정부에서는 매년 89,848건의 산업재해가 발생하고 있으며 2022년 현재까지 꾸준히 산업재해의 빈도수가 줄어들고는 있으나 여전히 많은 빈도수의 재해가 발생하고 있다. 2014년 산업재해보상보험법 적용사업장 17,062,308명 중에 4일 이상 요양을 요하는 재해자가 90,909명이 발생하였으며 직·간접비를 포함하여 19,632,795백만 원으로 큰 손실액이 발생하였다. 2021년에는 태안화력발전소에서의 재난 발생이후 정부에서는 '중대재해처벌법'을 도입하였고 산업재해를 줄이기 위해 기업과 개인에게 강조하고 있다. 따라서 산업재해로 겪는 각종 손실액과 인적 자원의 손실은 너무 크기 때문에 미리 조치를 통한 예측을 할 필요가 있다.

2017년 한국산업안전연구원에서 제공하는 산업재해 통계자료에 따르면 산업재해 빈도는 89,848건이 발생하였다. 건설업에서 28.55%로 25,649건의 가장 많은 건수의 산업재해가 발생하였다. 다음 제조업에서 28.2%로 25,333건의 재해자가 발생하였으며, 운수장고 통신업 4.72%(4,237명), 광업 2.11%(1,897명), 전기가스 수도업 0.1%(87명), 기타 36.33%(32,645명)이 발생하였다. 이에 가장 많은 비중을 차지하는 2개 업종 건설업과 제조업을 중심으로 마코프 체인 프로세스를 적용하여 산업재해 발생자에 대한 예측 모델을 제시하고자 한다. 그리고 실제 건설업과 제조업 내에서의 산업재해 발생값과 예측값을 비교하여 해당 모델링에 대한 증명을 통해 가까운 미래의 예측값을 제공하여 건설업과 제조업에 대한 산업재해 발생에 대한 확률 모형을 제시함으로써 산업재해 발생에 대한 예측에 도움을 주어 해당 업종에서의 유연한 대처를 위한 도움을 주고자 한다.

## 2. 선행연구

### 2.1 산업재해관련 연구

우리나라 산업재해 문제는 지속적으로 논의되어 왔으며 산업재해를 줄이기 위한 노력이 지속되고 있으며 이를 위해 다양한 예측 모델이 제안되고 있다. 오텔록 등은 XAI(eXplainable Artificial Intelligence)를

활용하여 산업재해 예측 모델링을 및 분석을 제안하였다[1]. 성현우 등은 통계적 자료를 통해서 건설 현장 사망사고 유형을 인자별로 분석하였다[2]. 강영식은 건설업에서 회귀분석을 적용하여 향후 5년간의 추세예측을 통해 건설업의 사망사고 예방을 제안하였다[3]. 김하영 등은 건설 현장의 복합 사고의 발생 패턴을 텍스트 마이닝을 이용해 분석하여 사고 유형중은 상대적으로 원인 또는 결과에 가깝게 작용하는 유형을 구분하였다[4]. 장길상 등은 제조업 현장에서의 안전관리를 위한 산업재해 예측 모델을 제안하고 제안된 모델을 기반으로 산업안전관리시스템(u-Safety 시스템을 제안하였다[5]. 홍성자 등은 중·소규모 제조업 사업장에서 발생하는 산업재해와 미신고된 산업재해를 평가하여 이들에 대한 지원정책과 정확한 전국적인 조사가 필요하다고 강조하였다[6]. 정병용과 이동하는 화학제품 제조업의 재해특성을 조사하여 사고 예방정책을 제안하였다[7].

산업재해 발생건수를 바탕으로 예측모델링을 제시하였으나 이들 데이터는 대부분 연간 단위에 대한 예측을 하였다는 점에서 업종별 업장에서의 세부적인 월별 대응에 둔감할거라는 문제가 발생할 것으로 사료된다.

### 2.2 마코프 체인 적용 연구

흔히 예측 모형으로 자주 사용되는 회귀모형은 회귀식을 통해 미래에 대한 예측이 가능하지만, 관찰된 사례가 부족하거나 일정한 패턴이 관찰되지 않을 경우 정확한 예측이 어렵다는 단점을 가진다. 반면 마코프 체인 프로세스는 과거추세의 시계열 규칙성이 없는 불특정한 패턴을 가진 자료 예측에 효과적이다. 손재환은 마코프 체인 프로세스에 기반으로 학교폭력 발생 예측 모델을 제시하였고[8], 정영석과 정진영은 전체 평균값, 1년 평균값, 최근 6개월 평균값을 적용하여 범죄 발생 예측 방법에 관한 방법론을 제시하였다[9]. 이희태, 김재철은 마코프 이론을 이용하여 전력 시스템 부품의 고장 횟수 예측 방법을 제시하였다[10]. 앞선 선행연구는 마코프 체인 프로세스는 발생 빈도가 불특정한 패턴의 자료에도 효과적으로 확률 예측 가능한 모형으로 알려져 있어 건설 산업재해발생에 적용하였다.

### 3. 연구방법

#### 3.1 마코프 체인 프로세스

마코프 체인을 이용한 건설업, 제조업 산업재해 발생 건수에 대한 예측 모델링을 구성해 보면 Fig. 1과 같다.

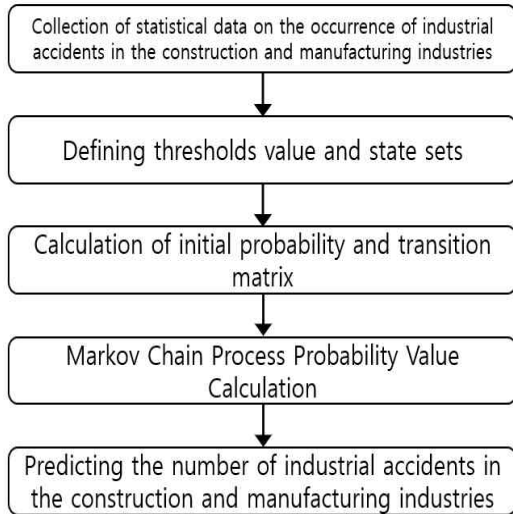


Fig. 1. Industrial accident prediction modeling using Markov chain process

본 연구에서 한국산업안전연구원에서 제공하는 산업재해 통계자료를 2014년 1월부터 2016년 12월까지 집계된 건설업, 제조업 재해 건수를 활용하였다.

Table 1. Number of reports of industrial accidents in the construction industry

	2014s'	2015s'	2016s'
Jan	2175	1649	2029
Feb	1396	1449	1501
Mar	1591	1515	1912
Apr	1945	1783	2026
May	1948	1838	2071
Jun	1969	2475	2438
Jul	2338	2430	2445
Aug	1999	2205	2586
Sep	1937	2419	2240
Oct	2166	2498	2271
Nov	2021	2354	2659
Dec	2184	2517	2392

Table 2. Number of reports of industrial accidents in the manufacturing industry

	2014s'	2015s'	2016s'
Jan	2493	2044	2099
Feb	2124	2064	1988
Mar	2352	2159	2379
Apr	2644	2360	2071
May	2360	2080	2103
Jun	2220	2594	2225
Jul	2730	2505	2357
Aug	2247	2026	2255
Sep	2237	2284	1921
Oct	2465	2240	2064
Nov	2258	21942	2383
Dec	2519	2461	2297

임계값 및 상태 집합 정의 단계에서 건설업, 제조업 산업재해 발생빈도의 상태를 정의한다. 상태 집합 단계에서는 건설 산업재해 빈도의 위험 상태가 가질 수 있는 범위를 나타낸다.

상태집합에서 정의한 위험 상태와 발생빈도 자료를 이용하여 위험 상태간의 전이상태를 확률로 나타낸 것이다. 각 건설업, 제조업 산업재해 빈도 데이터를 정의된 상태 집합과 매핑하여 상태들을 열거한다. 하나의 열거된 상태에서 다른 상태로의 전이행렬을 생성할수 있으며 상태전이 행렬 식(1)과 식(2)을 만족한다.

$$P = \begin{pmatrix} P_{11} & P_{12} & P_{13} & \cdots & P_{1n} \\ P_{21} & P_{22} & P_{23} & \cdots & P_{2n} \\ \cdots & \cdots & \cdots & P_{ij} & \cdots \\ P_{n1} & P_{n2} & P_{n3} & \cdots & P_{nn} \end{pmatrix} \quad (1)$$

$$\sum_{j=1}^n P_{1j} = 1, \sum_{j=1}^n P_{2j} = 1, \dots, \sum_{j=1}^n P_{nj} = 1, P_{ij} \geq 0$$

$$\sum_{j=1}^n P_{ij} = 1, i = 1, 2, \dots, n \quad (2)$$

정의된 각 위험 상태가 초기 상태에 발생 가능한 확률로 초기 확률값을 구한다. 식(3)은 초기 확률값을 위한 식으로 다음과 같다.

$$P(S_1, S_2, \dots, S_n) = P\left(\frac{a}{F}, \frac{b}{F}, \dots, \frac{c}{F}\right) \quad (3)$$

$$\sum_{i=1}^n P(S_i) = 1 \quad (4)$$

식(3)에서 a, b, c는 최근 빈도에서 정의된 각 상태 집합의 횟수이다. 최근 발생빈도에서 도출된 초기 확률값은 모두 더하면 1이 되어야 한다(식4). 본 논문에서는 2016년 8월부터 2016년 12월까지의 5개월 자료를 가지고 초기 확률값을 도출하였다.

발생 예측 단계에서는 전 단계에서 구한 위험 상태 전이행렬과 초기 확률값을 통해 앞으로 발생할 위험 발생 확률이나 빈도를 예측한다. 아래 식(5)는 건설업, 제조업 산업재해자 발생 확률식이다.

$$P(S_k) = \sum_{i=1}^n P(S_k)P_{ik} \quad (5)$$

식(5)에서 n은 재해 발생 상태 집합의 개수이고,  $P(S_i)$ 는 초기확률이며,  $P_{ik}$ 는 전이행렬이다.

## 4. 연구결과

### 4.1 건설업 산업재해 분석

#### 4.1.1 상태집합 정의

상태집합 범위의 발생 빈도를 분석해 위험 수준을  $S_1$ (관심상태),  $S_2$ (주의상태),  $S_3$ (경계상태)로 구분하여 정의하였다.  $S_1, S_2, S_3$ 을 구분하기 위해 평균으로부터 표준편차  $\pm 1$ 로 임계치를 설정하여 상태집단을 세 가지로 구분하였다. (평균=2093.639, 표준편차=344.4854, 범위 1396~2659, n=36, 최댓값= 2659). 표준편차 -1 미만에 해당되는 빈도들의 경우  $S_1$ (평균-표준편차 미만에 해당되는 값 : 2093-344=1749), -1과 +1 사이에 있는 빈도들  $S_2$ (평균-표준편차 이상 평균+표준편차 이하에 해당하는 값 : 2093-344=1749~2093+344=2439), 표준편차 +1 초과는 경계 상태  $S_3$ (평균+표준편차 초과에 해당하는 값: 2093+344=2439)로 정의하였다.

#### ◆임계값 범위

$S_1$ (관심상태): 1749미만

$S_2$ (주의상태): 1749이상 2439이하

$S_3$ (경계상태): 2439초과

#### 4.1.2 전이행렬 산출

Table 3은 2014년 1월부터 2016년 12월까지의 빈도를 위에서 정의한 건설업 산업재해 발생 건수를 임계값 범위에 매칭하여 상태를 나타낸다.

$S_2 S_1 S_1 S_2 S_2 S_2 S_2 S_2 S_2 S_2 S_2 S_2$

$S_1 S_1 S_1 S_2 S_2 S_3 S_2 S_2 S_3 S_2 S_3$

$S_2 S_1 S_2 S_2 S_2 S_2 S_3 S_3 S_2 S_2 S_3 S_2$

Table 3. Mapping frequency and state set of industrial accidents in the construction industry

	2014s'		2015s'		2016s'	
Jan	2175	$S_2$	1649	$S_1$	2029	$S_2$
Feb	1396	$S_1$	1449	$S_1$	1501	$S_1$
Mar	1591	$S_1$	1515	$S_1$	1912	$S_2$
Apr	1945	$S_2$	1783	$S_2$	2026	$S_2$
May	1948	$S_2$	1838	$S_2$	2071	$S_2$
Jun	1969	$S_2$	2475	$S_3$	2438	$S_2$
Jul	2338	$S_2$	2430	$S_2$	2445	$S_3$
Aug	1999	$S_2$	2205	$S_2$	2586	$S_3$
Sep	1937	$S_2$	2419	$S_2$	2240	$S_2$
Oct	2166	$S_2$	2498	$S_3$	2271	$S_2$
Nov	2021	$S_2$	2354	$S_2$	2659	$S_3$
Dec	2184	$S_2$	2517	$S_3$	2392	$S_2$

위 열거된 매칭된 값들로부터 각 상태에서 다른 상태로 전이되는 횟수의 빈도를 구하고 이를 바탕으로 상태전이 행렬식 (6)을 구하였다.

$$\begin{matrix}
 S_1 & S_2 & S_3 \\
 S_1 \begin{bmatrix} 3 & 3 & 0 \\ 3 & 15 & 5 \\ 0 & 5 & 1 \end{bmatrix} &
 S_1 \begin{bmatrix} 0.50 & 0.50 & 0.00 \\ 0.13 & 0.65 & 0.22 \\ 0.00 & 0.83 & 0.17 \end{bmatrix} &
 \end{matrix} \quad (6)$$

식 (6)을 다이어그램으로 나타내면 Fig. 2와 같다.

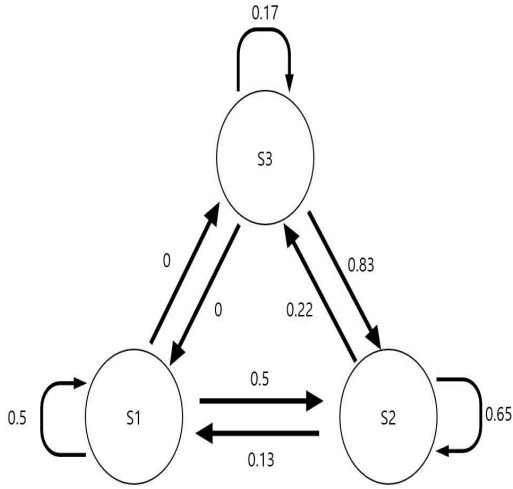


Fig. 2. State Diagram for Construction Occupational Accidents

4.1.3 초기확률 산출

건설업 산업재해 발생 예측 모델에 적용시키기 위한 초기 확률값을 산출하기 위해 2016년 5개월(2016년 8월~12월)간의 산업재해 발생건의 데이터를 사용하였다. Table 3을 이용하여 초기확률을 구하였다.

Table 4. Mapping frequency and state set of accidents in the construction industry for the last 5 months

Aug	Sep	Oct	Nov	Dec
2586	2240	2271	2659	2392
S <sub>3</sub>	S <sub>2</sub>	S <sub>2</sub>	S <sub>3</sub>	S <sub>2</sub>

Table 4에서 2016년 8월부터 12월까지의 상태집합을 살펴보면 S<sub>1</sub>:0개, S<sub>2</sub>:3개, S<sub>3</sub>:2개이다. 따라서 식(3)에 의해 도출되는 초기확률 값은 식(7)과 같다

$$P = (0, 0.6, 0.4) \tag{7}$$

4.1.4 발생예측

초기확률 식(7)과 상태전이 확률 식(6)을 이용하여 다음 달 건설업 산업재해 발생 확률을 구할 수 있다.

$$P(S_k) = \sum_{i=1}^n P(S_i)P_{ik} =$$

$$(0, 0.6, 0.4) \begin{bmatrix} 0.50 & 0.50 & 0.00 \\ 0.13 & 0.65 & 0.22 \\ 0.00 & 0.83 & 0.17 \end{bmatrix} \tag{8}$$

$$= (0.08, 0.72, 0.2)$$

구해진 식 8에 의해 다음 달 건설업 산업재해 발생 확률은 상태가 S<sub>2</sub>일 때 0.72로 가장 높게 나왔다. 다음 달에 발생할 건설업 산업재해 발생 건수를 예측하기 위하여 발생 확률 값에 건설업 산업재해의 전체 평균값 식(9), 1년 평균값 식(10), 단일 최대값 식(11)을 적용하였다.

◆건설업 산업재해 발생 예상 빈도:

$$\text{평균값 적용: } \sum_{i=1}^n P(S_i)M(S_i) \tag{9}$$

$$= 0.72 \times 2093 = 1517.13$$

$$\text{1년 평균값 적용: } \sum_{i=1}^n P(S_i)YM(S_i) \tag{10}$$

$$= 0.72 \times 2214 = 1604.47$$

$$\text{최대값 적용: } \sum_{i=1}^n P(S_i)Max(S_i) \tag{11}$$

$$= 0.72 \times 2659 = 1926.81$$

건설업 산업재해발생 건수를 예측한 결과 전체 평균값일때는 1518건, 1년 평균값일때는 1605건, 최대값을 적용했을때는 1927건 이었다.

4.2 제조업 산업재해 분석

4.2.1 상태집합 정의

건설업과 마찬가지로 상태집합 범위의 발생 빈도를 분석해 위험 수준을 S<sub>1</sub>(관심상태), S<sub>2</sub>(주의상태), S<sub>3</sub>(경계상태)로 구분하여 정의하였다. (평균=2272.278, 표준편차=194.9549, 범위 1921~2730, n=36, 최대값 = 2730). 표준편차 -1 미만에 해당하는 빈도들의 경우 S<sub>1</sub>(평균-표준편차 미만에 해당하는 값), -1과 +1 사이에 있는 빈도들 S<sub>2</sub>(평균-표준편차 이상 평균

+표준편차 이하에 해당하는 값), 표준편차 +1 초과는 경계 상태  $S_3$ (평균+표준편차 초과에 해당하는 값)로 정의하였다.

◆임계값 범위

- $S_1$ (관심상태): 2077미만
- $S_2$ (주의상태): 2077이상 2467이하
- $S_3$ (경계상태): 2467초과

4.2.2 전이행렬 산출

Table 5는 2014년 1월부터 2016년 12월까지의 빈도를 위에서 정의한 제조업 산업재해 발생건수를 임계값 범위에 매칭하여 상태를 나타낸다.

- $S_3 S_2 S_2 S_3 S_2 S_2 S_3 S_2 S_2 S_2 S_2 S_3$
- $S_1 S_1 S_2 S_2 S_2 S_3 S_3 S_1 S_1 S_2 S_2 S_2$
- $S_2 S_1 S_2 S_1 S_2 S_2 S_2 S_2 S_1 S_1 S_1 S_2 S_2$

Table 5. Mapping the frequency and state set of industrial accidents in the manufacturing industry

	2014s'		2015s'		2016s'	
<b>Jan</b>	2493	$S_3$	2044	$S_1$	2099	$S_2$
<b>Feb</b>	2124	$S_2$	2064	$S_1$	1988	$S_1$
<b>Mar</b>	2352	$S_2$	2159	$S_2$	2379	$S_2$
<b>Apr</b>	2644	$S_3$	2360	$S_2$	2071	$S_1$
<b>May</b>	2360	$S_2$	2080	$S_2$	2103	$S_2$
<b>Jun</b>	2220	$S_2$	2594	$S_3$	2225	$S_2$
<b>Jul</b>	2730	$S_3$	2505	$S_3$	2357	$S_2$
<b>Aug</b>	2247	$S_2$	2026	$S_1$	2255	$S_2$
<b>Sep</b>	2237	$S_2$	2284	$S_2$	1921	$S_1$
<b>Oct</b>	2465	$S_2$	2240	$S_2$	2064	$S_1$
<b>Nov</b>	2258	$S_2$	2194	$S_2$	2383	$S_2$
<b>Dec</b>	2519	$S_3$	2461	$S_2$	2297	$S_2$

위 열거된 매칭된 값들로부터 각 상태에서 다른 상태로 전이되는 횟수의 빈도를 구하고 이를 바탕으로 상태전이 행렬식 (12)을 구하였다.

$$\begin{matrix} S_1 & S_2 & S_3 \\ S_1 & \begin{bmatrix} 2 & 5 & 0 \\ 3 & 15 & 4 \\ 2 & 3 & 1 \end{bmatrix} \\ S_2 & \\ S_3 & \end{matrix} \quad \begin{matrix} S_1 & S_2 & S_3 \\ S_1 & \begin{bmatrix} 0.29 & 0.71 & 0.00 \\ 0.14 & 0.68 & 0.18 \\ 0.33 & 0.50 & 0.17 \end{bmatrix} \\ S_2 & \\ S_3 & \end{matrix} \quad (12)$$

식 (12)을 다이어그램으로 나타내면 Fig. 3와 같다.

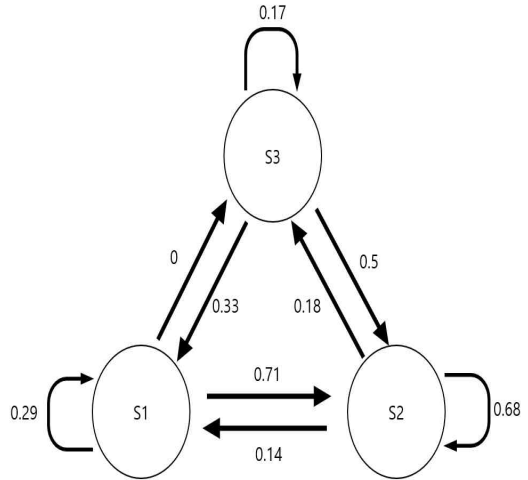


Fig. 3. State Diagram for Manufacturing Occupational Accidents

4.2.3 초기확률 산출

제조업 산업재해 발생 예측 모델에 적용시키기 위한 초기 확률값을 산출하기 위해 위해 최근 5개월 (2016년 8월~12월)간의 산업재해 발생건의 데이터를 사용하였다. Table 5를 이용하여 초기확률을 구하였다.

Table 6. Mapping of accident frequency and state set of accidents in the manufacturing industry for the past 5 months

Aug	Sep	Oct	Nov	Dec
2255	1921	2064	2383	2297
$S_2$	$S_1$	$S_1$	$S_2$	$S_2$

Table 6에서 2016년 8월부터 12월까지의 상태집합을 살펴보면  $S_1$ :2개,  $S_2$ :3개,  $S_3$ :0개이다. 따라서 식 (3)에 의해 도출되는 초기확률 값은 식(13)과 같다.

$$P = (0.4, 0.6, 0) \quad (13)$$

4.2.4 발생예측

초기확률 식(13)과 상태전이 확률 식(12)을 이용하여 다음 달 제조업 산업재해 발생 확률을 구할 수 있다.

$$P(S_k) = \sum_{i=1}^n P(S_i)P_{ik} =$$

$$(0.4, 0.6, 0) \begin{bmatrix} 0.29 & 0.71 & 0.00 \\ 0.14 & 0.68 & 0.18 \\ 0.33 & 0.50 & 0.17 \end{bmatrix} \quad (14)$$

$$= (0.20, 0.69, 0.11)$$

구해진 식 14에 의해 다음 달 제조업 산업재해 발생 확률은 상태가 S<sub>2</sub>일 때 0.69로 가장 높게 나왔다. 다음 달에 발생할 제조업 산업재해 발생 건수를 예측하기 위하여 발생 확률값에 제조업 산업재해의 전체 평균값 식(9), 1년 평균값 식(10), 단일 최댓값 식(11)을 적용하였다.

◆제조업 산업재해 발생 예상 빈도:

평균값적용:  $\sum_{i=1}^n P(S_i)M(S_i) \quad (9)$   
 $= 0.69 \times 2272 = 1578.79$

1년 평균값 적용:  $\sum_{i=1}^n P(S_i)YM(S_i) \quad (10)$   
 $= 0.69 \times 2178.5 = 1513.63$

최대값 적용:  $\sum_{i=1}^n P(S_i)Max(S_i) \quad (11)$   
 $= 0.69 \times 2730 = 1896.81$

제조업 산업재해발생 건수를 예측한 결과 전체 평균값일때는 1579건, 1년 평균값일때는 1514건, 최댓값을 적용했을때는 1897건 이었다.

4.3 평균값 및 최대값 적용 결과 분석

산업재해발생 빈도 건수중 가장 많은 유형을 차지하는 건설업과 제조업을 마코프 체인 프로세스에 적용하여 다음달 산업재해발생 확률을 산출하였다. 그리고 산업재해 발생 건수를 예측하기 위해 전체 평균값, 1년 평균값, 최댓값을 적용하여 Table 7에 나타내었다.

Table 7. Mapping frequency and state set of industrial accidents in the construction industry

		Predicted value	Actual value	Difference
Construction	Average	1 5 1 8	2 0 5 6	5 3 8
	Average of a year	1 6 0 5		4 5 1
	Maximum value	1 9 2 7		1 2 9
Manufacturing	Average	1 5 7 9	2 0 0 5	4 2 6
	Average of a year	1 5 1 4		4 8 1
	Maximum value	1 8 9 7		1 0 8

건설업은 전체 평균(2014년 1월~ 2016년 12월의 평균)을 적용하였을때 실제값(2017년 1월 데이터값)과 비교 하였을때 538의 차이가 났고, 1년 평균값(2016년 1월~2016년 12월의 평균)을 적용 하였을때 실제값과 451 차이가 났다. 최댓값을 적용 했을 때 129로 가장 작은 차이값을 보여주었다. 제조업의 경우 전체 평균(2014년 1월~ 2016년 12월의 평균)을 적용 하였을때 실제값(2017년 1월 데이터값)과 비교 하였을때 426의 차이가 났고, 1년 평균값(2016년 1월~2016년 12월의 평균)을 적용 하였을때 실제값과 481 차이가 났다. 최대값을 적용 했을때는 108로 가장 작은 차이값으로 실제 2017년 1월의 산업재해 발생건수와 가장 근접한 것을 보여준다.

## 5. 결론

최근 중대재해법에 대한 처벌이 강화되고 있으며 산업안전에 대해 더욱 강조되고 있다. 그에 따른 산업재해 예방에 대한 연구는 꾸준히 나오고 있으며 이에 대한 지속적인 노력이 필요하다.

따라서 본 논문에서는 산업재해 업종 중 가장 많은 비율을 차지하는 건설업과 제조업을 중심으로 산업재해 발생빈도에 대한 예측을 위한 예측 모형 모델을 제안하였다. 한국산업안전연구원에서 제공하는 2014년~2017년까지의 산업재해 자료를 바탕으로 마코프 체인 프로세스로 예측하였다. 마코프 체인 프로세스를 활용하여 다음달에 발생할 업종별 산업재해 발생 확률을 예측하였고 이를 전체 평균값, 1년 평균값, 최대값을 적용하여 산업재해 발생건수를 예측해 보았다. 해당 연구에 대한 결과로 최대값을 적용했을때 건설업에서 129의 가장 적은 차이, 제조업의 경우 108로 예측되었으며 2017년 1월의 실제 데이터값과 가장 근접함을 확인하였다. 따라서 본 연구에서 제안한 예측 모델링은 발생 빈도가 불특정한 특징을 가진 건설업과 제조업의 재해발생건수를 예측하기 위한 하나의 대안으로 적절하다고 생각한다.

본 논문에서 사용된 마코프 체인 프로세스를 적용하면 가까운 미래에 대한 산업재해 발생빈도를 예상할 수 있어 실제 산업재해 발생에 대한 발생 확률에 따라 경영책임자 및 현장책임자에 대한 실질적인 수치 제공과 더불어 전문인력 배치와 각종 시설물 및 기기·장치에 대한 안전점검과 수리를 하는등 현장에서 예방과 대응에 더욱 세심함을 줄것으로 예상된다. 또한 각종 관련 기관에서의 예방책에도 활용함으로써 구체적인 대응방법에 대해 모색할수 있도록 도움을 줄수 있을거라 예상된다. 더욱이 해당 연구에서의 건설업과 제조업뿐만 아니라 다른 업종에도 적용하여 보다 폭넓은 곳에서의 대처가 가능할거로 사료된다. 뿐만 아니라 각종 업종별 업무상 행동 및 사고 분석에도 적용하여 산업재해발생시 행동 유형별 위험 발생 요인을 분석하여 더욱 세심한 전략을 짤수 있을것로 예상된다.

그러나 마코프 체인 프로세스는 먼 미래를 예측하는게 힘들고, 산업재해는 예측할 수 없는 재난 사고라는 특성상 계절성을 가지지 않아 예측이 쉽지 않다는

한계점을 가지고 있다. 그래서 향후에는 과거 데이터로부터 현재 데이터까지의 모든 데이터를 바탕으로 딥러닝을 적용하여 더 정확한 미래에 예측을 할 수 있는 연구가 필요한 것으로 보인다.

## REFERENCES

- [1] H. S. OH, A. L. Son & Z. K. Lee. (2021). Title. Occupational accident prediction modeling and analysis using. *Journal of Digital Contents Society*, 22(7), 1115-1123.  
DOI : 10.9728/dcs.2021.22.7.1115
- [2] H. W. Seong, H. C. Choi, C. H. Roh, H. J. Son & C. W. Kim. (2016). Title. Critical Factors of fatality in Construction Industry : Statistical Analysis. *Critical Factors of fatality in Construction Industry : Statistical Analysis Conference*. 433-434
- [3] Y. S. Kang. (2017). Title. Department of Occupational Health & Safety Engineering, Semyung University. *J. Soc. Korea Ind. Syst.* 40(1), 87-94  
DOI : 10.11627/jkise.2017.40.1.087
- [4] H. Y. Kim, J. S. Yi & Y. E. Jang. (2022). Title. Analyzing Patterns of Multi-cause Accidents From KOSHA's Construction Injury Case Reports Utilizing Text Mining Methodology. *Journal of the Architectural Institute of Korea*, 38(4)  
DOI : 10.5659/JAIK.2022.38.4.237
- [5] G. S. Jang, & J. S. Lee. (2009). Implementation of u-Safety System for Preventing Industrial Disaster in Manufacturing Industries : The Case Study of 'H' Cooperation *The Journal of Information Systems*, 18(1), 145-157
- [6] S. J. Hong, M. J. Jeon & C. Y. Kim. (2011). The Actual State of Industrial Accidents in Small-medium Manufacturing Industries *Korean Journal of Occupational Health Nursing* 20(1), 93-103, May, 2011



- [7] B. Y. Jeong & D. H. Lee. (2001). Accident Characteristics and Prevention in Small Manufacturing Industries of Chemical Products *Journal of the Korea Safety Management & Science*. 3(2), 105-112, June, 2001
- [8] J. H. Son. (2015). A probabilistic model of school violence based on the markov chain process. *Studies on Korean Youth*. 26(2), 35-57  
DOI : 10.14816/sky.2015.26.2.35
- [9] S. Chung & J. J. Jung. (2012). A Study of the Prediction of Incidence of Crime using Markov proces. *Korean Society of Computer Information*. 17(3), 95-103
- [10] T. Lee & J. C. Kim. (2008). A Study on The Prediction of Number of Failures using Markov Chain and Fault Data. *Proceedings of the Korean Institute of Illuminating and Electrical Installation Engineers Conference*. 363-366
- [9]. S. Chung & J. J. Jung. (2012). A Study of the Prediction of Incidence of Crime using Markov proces. *Korean Society of Computer Information*. 17(3), 95-103
- [11] Ministry of Employment and Labor. (2014). 2014 Industrial Accident Status Analysis. Korea Occupational Safety Research Institute. Retrieved from <http://www.kosri.co.kr/board/index.html?id=board7&no=3>
- [12] Ministry of Employment and Labor. (2015). 2015 Industrial Accident Status Analysis. Korea Occupational Safety Research Institute. Retrieved from <http://www.kosri.co.kr/board/index.html?id=board7&no=5>
- [13] Ministry of Employment and Labor. (2016). 2016 Industrial Accident Status Analysis. Korea Occupational Safety Research Institute. Retrieved from <http://www.kosri.co.kr/board/index.html?id=board7&no=6>
- [14] Ministry of Employment and Labor. (2017). 2017 Industrial Accident Status Analysis. Korea Occupational Safety Research Institute. Retrieved from <http://www.kosri.co.kr/board/index.html?id=board7&no=9>
- [15] H. S. Lee. (2021). 2020 Status of Industrial Accidents. Korea Occupational Safety Research Institute. Retrieved from <http://www.kosri.co.kr/board/index.html?id=board7&no=13>

**박영진(Yeong Jin Park)****[학생회원]**

- 2016년 3월 ~ 현재 : 금오공과대학교 산업공학부 학사과정
- 관심분야 : 안전관리, 산업공학

· E-Mail : pyj5293@kumoh.ac.kr

**김태성(Tae Sung Kim)****[종신회원]**

- 2000년 12월 : Louisiana State University 산업공학과(박사)
- 2003년 3월 ~ 현재 : 금오공과대학교 산업공학부 교수
- 관심분야 : 최적화, Smart Factory, MES, SCM

· E-Mail : tkim@kumoh.ac.kr