

# 연관 규칙 탐색 기법을 이용한 건설공사 비사망 재해의 특성 요인 분석

이가연\* · 신성우\*\*†

## Analysis of Characteristic Factors for Non-fatal Accidents in Construction Projects using Association Rule Mining

Gayeon Lee\* · Sung Woo Shin\*\*†

### †Corresponding Author

Sung Woo Shin

Tel : +82-51-629-6473

E-mail : shinsw@pknu.ac.kr

Received : August 24, 2022

Revised : October 11, 2022

Accepted : October 20, 2022

**Abstract** : Simple statistical frequency based analysis, such as Pareto analysis, are widely used in conventional accident analysis. However, due to the dynamic and complex nature of construction works, many factors can simultaneously affect or involve the occurrence of accidents in construction projects. Therefore, the identification of the complex relationship between such factors is important to establish relevant and effective safety management policies and/or programs. In this study, characteristic factors and their relationships' contribution to non-fatal accidents in construction projects are analyzed using the association rule mining (ARM) technique. To this end, a total of 59,202 construction accident data are collected from 2015 to 2019 and the ARM is performed to retrieve specific relationships—named as association rules—among classified factors in the data. Characteristics of the retrieved relationships are analyzed and compared with the results of conventional Pareto analysis. Based on the results, it is found that both fall and trip are notable accident forms having characteristic relations with other factors for non-fatal accidents in construction projects. It is also found that small-scale construction, age of 50s, less than 1 month of working period, and architectural construction are important factors for non-fatal accidents in construction projects.

Copyright©2022 by The Korean Society of Safety All right reserved.

**Key Words** : accident analysis, association rule mining, non-fatal accident, construction safety

## 1. 서론

고용노동부가 2021년 발표한 2020년 산업재해 현황에 따르면 전체의 51.9%에 해당하는 458명의 사고사망자가 건설업에서 발생하였으며, 건설업의 사망만인율도 지난 2013년 이후 처음으로 2.0‰을 넘어 산재예방에 적신호가 켜진 상황이다<sup>1)</sup>. 제조업에 비해 건설업의 이러한 높은 산재 발생률은 건설공사의 작업자, 작업 조건, 작업 환경 등에 내재되어 있는 다양한 잠재 위험 요인들이 가변 복합적으로 작용하기 때문이다<sup>2)</sup>. 즉, 제조업에 비해 현장 상황 등에 따라 작업 조건이 수시로 변화하는 유동적 특성으로 인하여 건설업에서의 재해

또한 다양한 잠재 위험 요인들이 복합적으로 연계되어 발생한다<sup>3)</sup>. 따라서 건설업에 대한 효과적인 재해예방 정책이나 안전관리 대책의 수립을 위해서는 재해 분석 시에 이러한 복합 요인에 대한 분석이 필요하다<sup>4)</sup>. 그러나, 대부분의 산업재해분석 연구는 기인물 등에 대한 단순 기술 통계 분석이나 빈도가 높은 요인들 간의 교차분석이 주를 이루고 있고, 빈도가 낮은 요인들을 포함한 다양한 요인들이 동시에 발생하는 특성 등을 확률적이고 체계적으로 분석하는 연구는 일부 선도적 연구 외에는 아직 부족한 실정이다<sup>5-7)</sup>.

최근 AI 분야의 발전에 따라 데이터 마이닝(Data Mining) 등 다양한 머신러닝 기법을 산업재해분석에

\*국민건강보험공단 대구경북지역본부 주임 (Daegu-Gyeongbuk Regional Headquarter, National Health Insurance Service)

\*\*부경대학교 시스템경영·안전공학부 안전공학전공 교수 (Department of Safety Engineering, Pukyong National University)

적용하는 연구가 활발하게 이루어지고 있다<sup>4,8,9</sup>). 연관 규칙 탐색(ARM; Association Rule Mining)은 비지도 학습 기법의 일종으로, 어떤 사건에서 동시에 발생하는 변수들의 조합을 조건부 확률에 기반하여 도출할 수 있는 데이터 마이닝 기법이며, 대규모 데이터에서 변수 간의 잠재적인 연관 관계나 알려지지 않은 규칙을 밝혀낼 수 있는 방법이다. Weng et al.은 작업장에서 발생하는 충돌 재해에 대해 재해 발생에 영향을 미치는 요인을 ARM을 활용하여 분석하였다<sup>10</sup>). Cheng et al.은 대만에서 발생한 1,347건에 대한 데이터를 활용하여 건설업에서 일어나는 대부분 재해는 많은 잠재 요인으로 구성되어 있으며 이러한 요인들 간의 인과관계를 추론하기 위해 ARM을 활용하였다<sup>11</sup>). Son et al.은 낙하물에 기인한 건설 재해에 대한 연관 규칙을 분석하였고<sup>5</sup>), Lee et al.은 소규모 건설공사의 추락 사망 재해에 대해 공정률, 기인물, 직종, 연령, 요양기간, 근속기간, 재해발생시간, 시설물 등에 대한 변수를 활용하여 연관규칙을 도출하였다<sup>6</sup>). 또한 최근에 Lim et al.은 건설현장의 추락 사망자에 대한 연관 규칙을 분석하였다<sup>7</sup>). 기존의 연관규칙분석 연구들은 대부분 추락에 의한 사망 재해를 대상으로 하고 있으며, 추락 이외의 다른 형태의 재해나 사고성 산업재해의 대다수를 차지하는 비사망재해(Non-fatal Accidents)에 대한 연관규칙을 분석한 연구는 이루어지지 않은 한계가 있다.

본 연구에서는 ARM 기법을 적용하여 단순한 발생 빈도(frequency) 위주의 주요 요인 도출이 아닌, 발생 확률(probability)에 기반하여 건설공사의 재해에 영향을 미치는 여러 잠재 요인 중 동시에 자주 발생하는 공통 요인들 간의 발생 규칙에 대해 알아보하고자 한다. 특히, 추락 이외의 다른 형태의 재해나 비사망재해에 대해 분석하지 않은 기존 연구와는 달리, 본 연구에서는 여러 형태의 비사망재해에 대한 데이터를 바탕으로 건설 재해에 영향을 미치는 잠재 요인들 간의 동시 발생 규칙에 대해 알아보고, 이를 단순 빈도 위주의 통계 분석 결과와 비교하여 ARM 분석 결과에서 나타난 특징은 무엇인지를 알아보하고자 한다.

## 2. 연관 규칙 탐색 기법

ARM은 장바구니 분석(Market Basket Analysis) 기법으로도 알려진 비지도 데이터 마이닝 기법이며, 입력 데이터에 대한 결과를 If-Then의 형태로 반환하여 데이터의 If-Then 집합에서 숨겨진 규칙을 찾는 분석 기법이다. 여기서 If는 조건부(Antecedent)라고 하며, Then은 결과부(Consequent)라고 한다<sup>9</sup>). 만약 총  $N$ 개의 데이터

가 총  $m$ 개의 항목(또는 변수)으로 구성되어 있다면, 어떤 특정 데이터는 전체 항목 집합  $I = \{i_1, i_2, \dots, i_m\}$ 의 부분 집합이 된다. 연관 규칙( $R$ )은 조건부와 결과부로 구성되며, 어떤 항목 집합인  $X$ 와  $Y$ 에 대하여 ' $X$ (조건)가 일어나면  $Y$ (결과)도 일어난다'는 의미로 다음 식(1)과 같이 표현된다<sup>12</sup>).

$$R: X \Rightarrow Y \tag{1}$$

여기서,  $X, Y \subseteq I$ 이고  $X \neq Y$ 이다. 따라서, ARM은 적절한 항목 집합  $X, Y$ 를 찾는 문제로 볼 수 있다. ARM을 통해 찾아낸 규칙  $R$ 은 입력 데이터와 변수의 개수에 따라 수개에서 수천개가 발생할 수 있기 때문에, 일반적으로 지지도(Support), 신뢰도(Confidence), 향상도(Lift)의 세 가지 정량 척도를 이용하여 각 후보 규칙들에 대한 유용성을 평가하고, 유용성을 만족하는 규칙들을 의미있는 규칙으로 도출하는 전략을 취한다.

지지도( $S$ )는 전체  $N$ 개의 데이터 중에서  $X$ 와  $Y$ 가 동시에 나타나는 경우의 비율이며, 다음 식(2)와 같이 표현된다<sup>12</sup>).

$$S(R) = \frac{n(X, Y)}{N} = P(X \cap Y) \tag{2}$$

여기서,  $n(X, Y)$ 는 전체 데이터 중에서  $X$ 와  $Y$ 가 동시에 나타나는 경우의 개수이다. 식에서 알 수 있듯이 지지도는  $X$ 와  $Y$ 가 동시에 나타날 통계학적 확률을 의미하며, 교환법칙이 성립하기 때문에  $X \Rightarrow Y$ 의 지지도와  $Y \Rightarrow X$ 의 지지도는 같다.

신뢰도( $C$ )는 조건부 확률의 개념으로  $X$ 가 발생했을 경우에  $Y$ 도 동시에 발생할 확률을 의미하며, 다음 식(3)과 같이 표현된다<sup>12</sup>).

$$C(R) = P(Y|X) = \frac{P(X \cap Y)}{P(X)} = \frac{S(R)}{\frac{n(X)}{N}} \tag{3}$$

여기서,  $n(X)$ 는 전체 데이터 중에서  $X$ 가 포함된 데이터의 개수이다. 신뢰도는  $X$ 가 발생한 사건 중에서  $Y$ 도 동시에 발생한 사건의 비율이라고 해석할 수 있으며(예를 들어 신뢰도가 0.25일 경우,  $X$ 가 발생한 사건 4건 중  $X$ 와  $Y$ 가 동시에 발생한 사건은 1건이라는 의미), 따라서 신뢰도 값이 큰 규칙( $R$ )일수록 높은 의미를 갖는 규칙이라고 볼 수 있다.

마지막으로 향상도( $L$ )는 최소지지도와 최소신뢰도를 만족하는 규칙들 중 유용한 규칙을 필터링하거나

우선 순위를 결정하기 위해 사용하는 척도이며, 이는 결과부에 해당하는  $Y$ 가 단독으로 발생할 가능성에 비추어 조건과 연계하여 결과가 발생할 가능성(즉, 신뢰도)의 빈도 비율로 다음 식(4)와 같이 정의된다<sup>12)</sup>.

$$L(R) = \frac{C(R)}{n(Y)} = \frac{P(X \cap Y)}{P(X)P(Y)} \quad (4)$$

여기서,  $n(Y)$ 는 전체 데이터 중에서  $Y$ 가 포함된 데이터의 개수이다.  $P(X \cap Y) = P(X)P(Y)$ 인 경우는 조건부( $X$ )와 결과부( $Y$ )의 발생이 확률적으로 독립임을 나타내므로, 향상도가 1보다 큰 규칙은 결과( $Y$ )가 우연히 발생할 확률보다  $X$ 가 발생했을 때  $Y$ 가 발생할 확률이 더 높다는 것을 나타내며, 따라서 향상도 값이 1보다 크면 클수록 더 강한 연관성을 갖는 규칙이라고 평가할 수 있다<sup>12)</sup>.

### 3. 연구 방법

#### 3.1 연구 절차

본 연구는 Fig. 1과 같이 데이터 수집, 데이터 전처리, ARM 수행 및 결과 분석의 절차로 진행된다. 데이터 수집 및 전처리 단계에서는, 재해에 대한 다양한 정보가 포함된 건설 재해 데이터를 수집하고, 수집된 데이터 내의 세부 정보를 분석에 적합한 형태로 코드화를 실시한다. 다음으로 코드화된 데이터에 대한 Pareto 빈도 분석을 통해 상위 빈도를 차지하는 요인들을 도출하여 그 특성을 분석하고, 또한 ARM을 통해 확률론적으로 동시에 발생하는 요인들의 연관규칙을 도출하고 그 특성을 분석한다. 마지막으로 Pareto 분석 결과와 ARM 분석 결과를 비교하여 ARM 분석의 특징에 대해 알아본다.

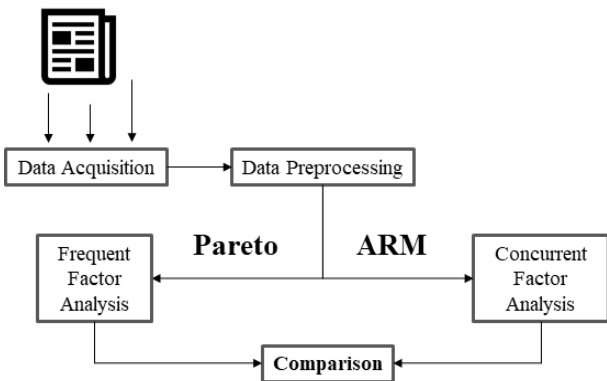


Fig. 1. Schematic Research Procedures.

#### 3.2 데이터 수집 및 전처리

본 연구에서는 지난 2015년부터 2019년까지 5년 동안 고용노동부에 보고된 115,904건의 비사망 건설 재해 데이터를 수집하였다. 본 연구는 사고성 재해에 대한 분석을 목적으로 하였기 때문에 질병 관련 재해자에 대한 데이터 49,925건은 제외하였으며, 또한 남은 데이터 중에서 다음에 설명할 코드화 과정에서 각 항목 코드별 정보가 없거나 분류 불능이 포함된 6,777건의 데이터도 분석 대상 데이터에서 제외하였다. 이에 따라 최종적으로 총 59,202건의 데이터를 분석 대상 데이터로 선택하였다.

한편, 수집된 데이터에 제시된 정보는 Table 1과 같이, 공사와 관련된 정보, 현장과 관련된 정보, 재해부상자와 관련된 정보, 재해와 관련된 정보를 포함하고 있다. 본 연구에서는 이러한 여러 정보들 중에서 재해에 영향을 미치는 주요 잠재 요인으로 기존 연구에서 공통되게 많이 선택한 공사종류(Construction Types: CT), 공사진행률(Progress Rates: PR), 공사규모(Construction Scales: CS), 재해형태(Accident Forms: AF), 기인물(Original Cause Materials: OCM), 직종(Occupations: O), 연령(Ages: A), 동종업무 근속기간(Working Years: WY), 재해발생시간(Accident Times: AT), 상해요양기간(Medical Treatment Times: MTP), 공사시설물종류(Facility Types: F)을 분석 대상 변수로 선정하였다<sup>5-7,11)</sup>.

Table 1. Types of information contained in research data

Category	Types of information	Selection
Project information	Contractor information	-
	Site address	-
	Administrative district	-
Site information	Construction types (Large, Medium, Small)	O
	Progress rate	O
	Construction period	-
	Number of workers	-
	Construction scale	O
	Facility types (Small, Medium, Large)	O
Injured worker information	Personal information (Name, Age, Sex, etc.)	O
	Occupation (Large, Medium, Small)	O
	Working years	O
	Employment type	-
Accident information	Accident day	-
	Accident time	O
	Accident forms (Large, Medium, Detail)	O
	Original cause material (Small, Medium, Large, Detail)	O
	Medical treatment period	O
	Date of insurance approval	-

Table 2. Classification of variables and factors

Variables	Factors (Sub-variables)
Occupations (15)	O-1 (Skilled building workers), O-2 (Non-skilled simple laborers), O-3 (Earthwork workers), O-4 (Earthwork machinery Operators), O-5 (Skilled finishing workers), O-6 (Machinery related workers), O-7 (Electrical workers), O-8 (Painters), O-9 (Pipelayers), O-10 (Welders), O-11 (Vehicle drivers), O-12 (Gardening workers), O-13 (Other skilled workers), O-14 (Field Engineers), O-15 (Field managers)
Accident times (6)	AT-1 (07:00~10:00), AT-2 (10-13), AT-3 (13~16), AT-4 (16~19), AT-5 (19~22), AT-6 (22~07)
Medical treatment periods (7)	MTP-1 (More than 24 months), MTP-2 (12~24 months), MTP-3 (6~12 months), MTP-4 (3~6 months), MTP-5 (1~3 months), MTP-6 (2 weeks~1 month), MTP-7 (Less than 2 weeks)
Ages (6)	A-1 (10~19 years), A-2 (20~29), A-3 (30~39), A-4 (40~49), A-5 (50~59), A-6 (More than 60)
Facilities (25)	F-1 (Single house), F-2 (Apartment), F-3 (Water Supply), F-4 (Transportation), F-5 (Neighborhood), F-6 (Environmental), F-7 (Heavy industrial), F-8 (Light industrial), F-9 (Electric/Electronic), F-10 (Petrochemical), F-11 (Sports/Exhibition), F-12 (Education/Government/Research), F-13 (Agricultural), F-14 (Harbor), F-15 (Medical/Office), F-16 (Security), F-17 (Energy Supply), F-18 (Resource Storage), F-19 (Information Communication), F-20 (Bridge), F-21 (Tunnel), F-22 (Landscape), F-23 (Historical Architect), F-24 (Complex Site), F-25 (Others)
Construction types (7)	CT-1 (Architectural construction), CT-2 (Dam/Embankment construction), CT-3 (Machine/Facility installation construction), CT-4 (Road construction), CT-5 (Railroad construction), CT-6 (Hydropower construction), CT-7 (Others)
Accident forms (15)	AF-1 (Fall), AF-2(Trip), AF-3 (Struck-by), AF-4 (Cut), AF-5 (Knock), AF-6 (Caught-in), AF-7 (Overturn), AF-8 (Collapse), AF-9 (Traffic Accident), AF-10 (Electric Shock), AF-11 (Fire/Explosion), AF-12 (Abnormal temperature contact), AF-13 (Chemical materials contact/Exposure), AF-14 (Drowning), AF-15 (Choke)
Original cause materials (39)	OCM-1 (Ladders), OCM-2 (Floors/Aisles), OCM-3 (Scaffolds), OCM-4 (Powered hand equipments), OCM-5 (Work plates), OCM-6 (Forms), OCM-7 (Metallic materials), OCM-8 (Processing/Cutting machinery), OCM-9 (Non-metallic materials), OCM-10 (Hand equipments), OCM-11 (Stairs), OCM-12 (End/Openings), OCM-13 (Building surfaces), OCM-14 (Special Vehicles), OCM-15 (Woods), OCM-16 (Parts/Appendages), OCM-17 (Roofs/Girders), OCM-18 (Excavation/Loading Machinery), OCM-19 (Slopes/Rocks), OCM-20 (Lifting machinery), OCM-21 (Steel Beams/Trusses), OCM-22 (Temporary Walkways/Guardrails), OCM-23 (Building components), OCM-24 (Power Transmission equipments), OCM-25 (Form Supports), OCM-26 (Manholes/Tunnels), OCM-27 (Water/Vapor), OCM-28 (Mechanical equipments), OCM-29 (Containing/Packaging Tools), OCM-30 (Electrical equipments), OCM-31 (Compaction/Paving machinery), OCM-32 (Groundwork machinery), OCM-33 (Chemical materials), OCM-34 (Tunneling machinery), OCM-35 (Porklifts), OCM-36 (Temporary earth-retainers), OCM-37 (Heaters), OCM-38 (Other construction machinery), OCM-39 (Miscellaneous)
Working years (7)	WY-1 (less than 1 month), WY-2 (1~2 months), WY-3 (2~3 months), WY-4 (3~6 months), WY-5 (6~12 months), WY-6 (12~24 months), WY-7 (More than 24 months)
Process rates (6)	PR-1 (0~20%), PR-2 (20~40%), PR-3 (40~60%), PR-4 (60~80%), PR-5 (80~100%), PR-6 (After Completion)
Construction scales (4)	CS-1 (below 0.3 billion), CS-2 (0.3~5.0 billion), CS-3 (5.0~12.0 billion), CS-4 (over 12.0 billion)

선정된 11 가지 분석 대상 변수의 세부 정보를 코드화하기 위해, 각 변수별 하위 요소를 세부 변수(항목)으로 분류하고 이를 코드화하는 전처리를 실시하였다. 세부 항목 분류는 ‘산업재해 기록분류에 관한 지침 (KOSHA Guide G-83-2016)’에 제시된 분류를 기준으로 하되, 지나친 세분화가 되지 않도록 일부 유사한 항목들은 하나의 항목으로 통합하거나 범위를 조정하여 처리하였다. 본 연구에서 분류한 분석 대상 변수별 세부 항목 코드는 Table 2와 같다. 직종은 건설관련기능인력 (O-1)을 포함한 총 15개의 항목으로 코드화하였으며, 재해발생시간(AT)은 6개의 세부 항목으로, 재해의 심각도를 알 수 있는 상해요양기간은 7개의 세부 항목으로 분류하였다. 또한, 공사시설물종류, 재해발생형태, 기인물은 각각 25개, 15개, 39개의 세부 항목으로 분류하였으며, 재해자의 연령대, 근속기간, 공사종류, 공사진행률, 공사규모는 각각 6개, 7개, 7개, 6개, 4개의 세부 항목으로 분류하였다.

### 3.3 연관규칙탐색

ARM은 일반적으로 지지도, 신뢰도, 향상도의 세 가지 척도를 이용하여 데이터 속에서 동시에 발생하는 항목(요인)들 간의 발생 규칙을 찾아내기 위한 기법이며, 이를 컴퓨터를 이용하여 효과적으로 수행할 수 있는 여러 알고리즘이 개발되었다. 본 연구에서는 기존에 개발된 알고리즘 중에서 효율성이 높아 ARM에서 보편적으로 많이 사용되는 APRIORI 알고리즘을 적용하였다. APRIORI 알고리즘은 미리 설정한 최소 지지도 이상의 지지도를 갖는 빈발항목집합을 데이터로부터 생성하고, 생성된 빈발항목집합의 각 부분집합 중에서 미리 설정한 최소 신뢰도 이상의 신뢰도를 갖는 부분집합을 찾아 향상도가 높은 순으로  $R: X \Rightarrow Y$  형태의 규칙을 반환한다<sup>12)</sup>. APRIORI 알고리즘은 다양한 소프트웨어 프로그램에 Package나 Library 형태로 구현되어 있으며, 본 연구에서는 Python™ 프로그램에서 제공하는 Package를 이용하였다.

APRIORI 알고리즘을 이용한 ARM 시에 유의할 점은 최소 지지도의 설정에 있다. 만약 최소 지지도를 너무 높게 설정할 경우에는 전체 데이터 중에서  $X$ 와  $Y$ 가 동시에 존재하는 빈발항목집합의 크기가 매우 작아지며, 이에 따라 동시에 나타나는 빈도는 상대적으로 낮지만 신뢰도나 향상도는 큰 규칙은 누락될 수 있다. 예를 들어, 최소 지지도가 0.2로 설정된 경우, APRIORI 알고리즘에서는 전체 데이터 중에서  $X$ 와  $Y$ 가 동시에 존재하는 데이터가 적어도 5개 중에 1개 이상이어야 이를 후보(즉, 빈발항목집합)로 생성하게 되며, 0.2보다 작은(예를 들어 10개 중에 1개) 후보는 비록 이 후보가 다른 후보에 비해서 신뢰도나 향상도가 크더라도 빈발항목 집합에서 제외되게 된다. 한편, 최소 지지도의 값을 너무 낮출 경우에는 빈도가 낮더라도 신뢰도나 향상도가 큰 규칙이 누락되지 않을 수 있으나, 데이터 규모에 따라 의미가 매우 낮은 규칙이 도출될 수 있는 단점도 존재한다. 예를 들어, 데이터의 규모가 100건인 경우에 최소 지지도를 0.01로 설정하면, 전체 데이터 100건 중에 빈도가 1건인 경우도 규칙으로 도출될 수 있다. 따라서, 최소 지지도는 전체 데이터의 규모를 고려하여 결정할 필요가 있다. 본 연구에서는 전체 데이터 규모가 59,202건으로 많다는 점을 고려하여 최소 지지도를 0.015(즉, 약 900건 이상의 동시 발생 빈도를 나타내는 경우)로 설정하였다.

한편, 지지도는 교환법칙이 성립하기 때문에  $X \Rightarrow Y$ 의 지지도와  $Y \Rightarrow X$ 의 지지도는 같다. 따라서, 같은 지지도 값을 갖는 경우,  $X \Rightarrow Y$  형태의 규칙인지,  $Y \Rightarrow X$  형태의 규칙인지를 판정하기 위해서는 신뢰도 값을 비교하여야 한다. 최소 신뢰도는 이를 판정하기 위한 최소 기준값을 의미하며, 주로 0.5를 기준값으로 많이 사용한다<sup>12)</sup>. 이에 따라, 본 연구에서도 최소 신뢰도 값을 0.5로 설정하고 ARM을 실시하였다.

## 4. 결과 및 토의

### 4.1 Pareto 분석

전체 데이터의 경향성을 파악하기 위해 단순 빈도 위주의 Pareto 분석을 실시하였다. Table 3은 전체 데이터에서 출현 빈도수가 10% 이상인 항목을 나타낸다. 가장 빈도수가 높은 항목은 근속기간 1개월 미만(WY-1)으로, 전체 59,202건의 데이터 중에 41,402건의 데이터에 WY-1이 포함된 것으로 나타났다. 근속기간 1개월~2개월(WY-2)은 21번째로 빈도수가 높은 항목으로 7131건의 데이터에 포함된 것으로 나타났다. 이는 전체 재해(59,202건) 중 48,533건(41,402건+7,131건)이

근속기간 2개월 이내에 발생하였음을 나타낸다. 각 변수별로 상위 빈도수를 차지하는 항목을 보면, CT는 CT-1이, CS는 CS-2, A는 A-5, MTP는 MTP-4, O는 O-1, AF는 AF-1이, PR은 PR-1, F는 F-1, AT는 AT-3이 가장 높은 빈도를 차지하는 항목들로 나타났다.

한편, Table 3에서 빈도수가 10% 이상인 항목 중에 기인물(OCM) 변수에 대한 항목은 나타나지 않았는데, 이는 다른 변수에 비해 기인물 변수에 속하는 세부 항목의 개수(39개)가 많고 또한 특정 항목이 전체를 지배하지 않고 비중이 여러 항목에 고르게 분포하기 때문으로 판단된다. Table 4는 기인물에 대한 Pareto 분석 결과를 나타낸다. 총 39개의 기인물 중에서 13개의 기인물이 전체의 75% 이상을 차지하고 있으나, 가장 높은 빈도를 나타내는 사다리(8.5%)와 표 내에서 가장 낮은 빈도를 차지하는 건축물·구조물 표면(3.1%)의 비율 차이가 크지 않고 상대적으로 고른 비율 분포를 나타내고 있음을 알 수 있다.

Pareto 분석 결과를 살펴 보면, 건설 재해는 (1) 공사 금액이 작은 소규모 건설공사에서, (2) 연령대는 50대에서, (3) 공사종류는 건축공사에서, (4) 재해유형은 떨어짐이 지배적으로 발생한다는 기존의 재해 통계 분석 결과와 잘 일치하고 있음을 알 수 있다<sup>13,14)</sup>. 그러나,

Table 3. Pareto analysis for overall data

Rank	Factors	Cases (%)	Rank	Factors	Cases (%)
1	WY-1	41402 (69.93)	12	MTP-5	17565 (29.67)
2	CT-1	40265 (68.01)	13	A-4	14837 (25.06)
3	CS-2	26096 (44.07)	14	F-1	13312 (22.49)
4	A-5	25197 (42.56)	15	PR-2	12752 (21.54)
5	CS-1	24057 (40.64)	16	MTP-3	11610 (19.61)
6	MTP-4	23379 (39.49)	17	PR-1	10074 (17.02)
7	O-1	22504 (38.01)	18	PR-5	9769 (16.50)
8	AF-1	20215 (34.15)	19	AF-2	9678 (16.50)
9	O-2	18658 (31.52)	20	AF-3	8449 (14.27)
10	CT-7	18238 (30.80)	21	WY-2	7131 (12.45)
11	AT-3	18142 (30.64)	22	F-2	6837 (11.55)

Table 4. Pareto analysis for original cause materials

Rank	Factors	Cases (%)	Rank	Factors	Cases (%)
1	OCM-1	5047 (8.53)	8	OCM-8	2592 (4.38)
2	OCM-2	4956 (8.37)	9	OCM-9	2458 (4.15)
3	OCM-3	4701 (7.94)	10	OCM-10	2299 (3.88)
4	OCM-4	4694 (7.93)	11	OCM-11	2016 (3.41)
5	OCM-5	4424 (7.47)	12	OCM-12	1866 (3.15)
6	OCM-6	4075 (6.88)	13	OCM-13	1833 (3.10)
7	OCM-7	3787 (6.40)	Total		44748 (75.59)



Table 5. Pareto analysis for accident forms

Rank	Factors	Cases (%)	Rank	Factors	Cases (%)
1	AF-1	20215(34.15)	9	AF-9	538 (0.91)
2	AF-2	9678 (16.35)	10	AF-10	471 (0.80)
3	AF-3	8449 (14.27)	11	AF-11	274 (0.46)
4	AF-4	6046 (10.21)	12	AF-12	159 (0.27)
5	AF-5	5189 (8.76)	13	AF-13	71 (0.12)
6	AF-6	5080 (8.58)	14	AF-14	33 (0.06)
7	AF-7	1983 (3.35)	15	AF-15	22 (0.04)
8	AF-8	994 (1.68)	Total		59202 (100)

이러한 Pareto 분석과 같은 단순 빈도 위주의 통계 분석은 어느 한 특정 변수의 세부 항목들 간의 발생 가능성을 분석하기에는 용이하나, 여러 변수들의 각 세부 항목 상호 간에 존재하는 규칙이나 패턴을 정량적으로 파악하기는 어렵다. 예를 들어, Table 5는 재해발생형태(AF)에 대한 Pareto 분석 결과를 나타내며, 가장 빈도가 높은 (즉, 발생 가능성이 가장 높은) 재해형태는 떨어짐임을 알 수 있다. 그러나, Table 4와 Table 5를 조합하여 “사다리(X)에서 떨어짐(Y)이 발생할 가능성이 높을 것”이라는 추론을 제시할 수는 있으나, (1) 이러한 가능성이 정량적으로 어느 정도(즉,  $P(X \cap Y)$ )이며, 다른 가능성(예를 들어, 비계에서 떨어짐)에 비해서는 어떠한지는 알기 어렵다. 뿐만 아니라, (2) ‘떨어짐’ 재해가 발생하였을 때 그 기인물이 ‘사다리’일 가능성이 높은지(즉,  $X \Rightarrow Y$ ) 또는 기인물이 ‘사다리’일 때 ‘떨어짐’ 재해가 발생할 가능성이 높은지(즉,  $Y \Rightarrow X$ )도 알 수가 없다. 더 나아가, Table 3과 조합하여, “근속기간이 1개월 미만(WY-1)인 근로자가, 공사금액이 3억원~50억원(CS-2)인, 건축공사(CT-1) 현장에서, 떨어짐 재해(AF-1)가 발생할 가능성”( $\{WY-1, CS-2, CT-1\} \Rightarrow \{AF-1\}$ )이 있음을 추론할 수는 있으나, 이러한 가능성이 “동일한 조건에서 넘어짐 재해(AF-2)가 발생할 가능성( $\{WY-1, CS-2, CT-1\} \Rightarrow \{AF-2\}$ )”에 비해 더 높은 가능성을 가지는지에 대해서도 판단할 수가 없다. 따라서, 여러 변수들의 각 세부 항목 상호 간의 이러한 다양한 조합에 대한 가능성을 정량적으로 파악하고, 가능성이 높은 조합들 중에서도 어느 조합이 다른 조합에 비해 더 높은 가능성을 갖는가를 파악하기 위해서는 ARM 분석이 필요하다.

#### 4.2 ARM 분석

전체 데이터 59,202건에 대해 최소 지지도 0.015와 최소 신뢰도 0.5로 설정하여 ARM을 실시하였고, 이를 통해 총 637개의 규칙을 구하였다. 구해진 규칙의 향상

도 값은 모두 1.0 이상의 값(최소 1.501, 최대 7.881)으로 나타나, 강한 연관성이 있는 규칙들이 구해졌음을 알 수 있다.

Table 6은 ARM 분석을 통해 구해진 총 637개의 규칙들 중에서 각각 지지도(S), 향상도(L), 신뢰도(C)를 기준으로 상위에 해당하는 규칙의 일부를 나타내었다. 지지도 값이 가장 큰 규칙은, 조건이 “단독주택(F-1) 공사에서 근속기간 1개월 미만(WY-1)”일 때 “공사 규모 3억원 이하(CS-1) 건축공사(CT-1)이다”로 나타났다. 향상도와 신뢰도가 가장 높은 규칙은, 각각 조건이 “기인물이 동력형 휴대용공구(OCM-4)”일 경우, “건축공사(CT-1)에서 절림·베임·절단(AF-4)” 재해가 발생할 가능성이 높다는 규칙과 “근속기간 1개월 미만(WY-1)인

Table 6. Upper ranked association rules

Rank Ordered by support (S)					
Rank	Association rules (R)		S	C	L
	If (X)	Then (Y)			
1	{WY-1, F-1}	{CT-1, CS-1}	0.092	0.538	2.299
2	{OCM-2}	{AF-2}	0.082	0.982	6.009
3	{AF-2}	{OCM-2}	0.082	0.503	6.009
4	{OCM-1}	{AF-1}	0.081	0.946	2.771
5	{OCM-3}	{AF-1}	0.066	0.834	2.442
6	{OCM-5}	{AF-1}	0.060	0.800	2.342
Rank Ordered by Lift (L)					
Rank	Association rules (R)		S	C	L
	If (X)	Then (Y)			
1	{OCM-4}	{AF-4, CT-1}	0.017	0.536	7.881
2	{F-1, OCM-2}	{AF-2, CT-1}	0.016	0.894	7.775
3	{OCM-4, CS-1}	{AF-4, WY-1}	0.023	0.551	7.409
4	{OCM-4, WY-1, O-1}	{AF-4}	0.017	0.746	7.301
5	{OCM-4, O-1}	{AF-4}	0.024	0.739	7.241
6	{AF-4, CS-1}	{OCM-4}	0.030	0.574	7.238
Rank Ordered by Confidence (C)					
Rank	Association rules (R)		S	C	L
	If (X)	Then (Y)			
1	{OCM-2, CT-1, O-1, WY-1}	{AF-2}	0.018	0.991	6.065
2	{OCM-2, CT-1, CS-2, WY-1}	{AF-2}	0.020	0.991	6.060
3	{OCM-2, WY-1, AT-3}	{AF-2}	0.017	0.990	6.058
4	{OCM-2, O-1, WY-1}	{AF-2}	0.023	0.990	6.055
5	{OCM-2, CT-1, AT-3}	{AF-2}	0.017	0.989	6.051
6	{OCM-2, A-5, CS-2}	{AF-2}	0.017	0.989	6.051

건설관련기능인력(O-1)이 건축공사(CT-1) 현장에서 바닥·통로(OCM-2)가 기인물인” 경우 “넘어짐(AF-2)” 재해가 발생할 가능성이 높다는 규칙으로 나타났다. 지지도가 상위인 규칙들은 앞서 Pareto 분석에서 가장 상위에 있는 항목들 간의 단순한 형태(예를 들어, 기인물이 사다리(OCM-1)이면 떨어짐(AF-1) 재해가 발생할 가능성이 높다) 규칙들이 상위 규칙으로 많이 나타났으며, 특히 “공사규모 3억 미만(CS-1), 근속기간 1개월 미만(WY-1), 건설관련기능인력(O-1), 건축공사(CT-1), 떨어짐(AF-1), 넘어짐(AF-2), 단독주택공사(F-1), 사다리(OCM-1), 비계(OCM-3), 작업발판(OCM-5)”를 항목으로 하는 규칙이 많이 나타났다.

반면에 Table 6에서 알 수 있듯이 향상도나 신뢰도가 높은 규칙들은 Pareto 분석에서 상위에 있는 항목 뿐만 아니라 하위에 해당하는 항목들도 포함하여 좀 더 복합적인 형태(즉, 조건절이나 결과절에 포함된 항목이 다수인)의 규칙들이 상위 규칙에 많이 나타났다. 예를 들어, 향상도 상위 규칙에는 “동력형 휴대형공구(OCM-4), 찰림·베임·절단(AF-4), 바닥·통로(OCM-2), 넘어짐(AF-2), 요양기간 3~6개월(MTP-4), 단순노무자(O-2)” 등 지지도 기준 상위 규칙에 포함되지 않은 항목들로 구성된 규칙이 상위 규칙으로 많이 나타났다. 마지막으로 신뢰도 기준 상위 규칙에는 지지도 기준 상위 규칙과 향상도 기준 상위 규칙에 포함된 항목들 중 재해형태를 제외한 항목들이 조건절에 복합적으로 나타난 규칙들이 상위 규칙에 많이 나타났으며, 상위 규칙들의 결과절은 대부분 재해형태 항목인 “떨어짐(AF-1), 넘어짐(AF-2), 찰림·베임·절단(AF-4)”이 포함된 형태로 나타났다.

주목할만한 점은 향상도 기준 상위 100위 이내에는 떨어짐(AF-1)과 이와 관련된 기인물 항목(예를 들어, 사다리, 비계, 작업발판 등)들이 포함된 규칙이 나타나지 않았다는 점과 신뢰도 기준 100위 이내에는 Pareto 분석 기준 4위에 해당하는 “찰림·베임·절단(AF-4)” 항목이 포함된 규칙은 있으나, 3위에 해당하는 “물체에 맞음(AF-3)” 항목이 포함된 규칙은 이보다 하위 규칙에서 나타났다는 점이다. 이는 지지도가 여러 항목들 간의 조합 중에서 가장 높은 빈도를 갖는 조합을 찾는 지표인 반면에, 신뢰도와 향상도는 규칙의 형태적인 측면에서 더 강한 연관성을 나타내는 규칙을 찾는 지표이기 때문으로 사료된다. 즉, 향상도는 결과절의 단독 발생 확률에 비해 조건절이 주어진 경우에 결과절의 발생 확률이 더 높은가를 나타내는 지표이기 때문에, (1) 조건절이 동일한 경우 결과절에 나타난 항목의 데이터 내 발생 빈도가 상대적 높은 경우나 (2) 결

과절이 동일할 때 조건절에 나타난 항목의 데이터 내 발생 빈도가 상대적으로 높은 경우에는 향상도 값은 낮아지게 된다. 따라서, 향상도는 조건절이 동일하고 결과절이 다른 규칙들이나, 결과절이 동일하고 조건절이 다른 규칙들 중에서 더 높은 가능성을 갖는 규칙을 판단하고자 할 경우(예를 들어 Table 6에서 지지도 기준 4위, 5위, 6위의 규칙)에 특히 유용성이 높다. 한편, 신뢰도는 일정한 향상도 이상의 값을 갖는 규칙들 중에서 다른 규칙들에 비해 주어진 조건절과 결과절 간의 인과성이 더 높은 규칙을 판단하고자 할 경우에 특히 유용성이 높다. 예를 들어, 지지도 기준 2위와 3위의 규칙과 같이 조건절과 결과절이 서로 바뀐 상호 대응형 규칙은, 식(2)와 식(4)에서 알 수 있듯이 지지도와 향상도는 교환법칙의 성립에 따라 같은 값을 갖지만 신뢰도는 다른 값을 가지며, 이 경우 신뢰도가 높은 규칙이 더 높은 인과성을 갖는 규칙이라고 볼 수 있다. 본 연구에서 신뢰도 기준 상위 300위 까지의 규칙들은 “기인물(OCM)⇒재해형태(AF)”와 같이 인과적 형태의 규칙들이 대부분(271개)을 차지하였는데, 이는 신뢰도 지표의 이러한 인과적 특성을 보여주는 결과이다.

Table 7. Characteristic association rules for non-fatal accidents in construction projects

Association rules (R)		S	C	L
If (X)	Then (Y)			
{AT-3, OCM-2}	{AF-2}	0.025	0.989	6.049
{O-1, OCM-2}	{AF-2}	0.033	0.988	6.045
{CT-1, OCM-2, WY-1}	{AF-2}	0.039	0.986	6.033
{A-5, OCM-2, WY-1}	{AF-2}	0.025	0.985	6.028
{OCM-2, WY-1}	{AF-2}	0.057	0.985	6.023
{MTP-5, OCM-2}	{AF-2}	0.028	0.982	6.006
{OCM-2}	{AF-2}	0.082	0.982	6.009
{O-2, OCM-2}	{AF-2}	0.028	0.979	5.989
{CT-7, OCM-2}	{AF-2}	0.026	0.977	5.977
{CS-1, OCM-1}	{AF-1}	0.046	0.957	2.872
{MTP-4, OCM-1}	{AF-1}	0.034	0.952	2.788
{CT-7, OCM-1}	{AF-1}	0.032	0.947	2.774
{OCM-1}	{AF-1}	0.081	0.946	2.771
{AT-3, OCM-1}	{AF-1}	0.027	0.945	2.767
{CT-1, OCM-1, WY-1}	{AF-1}	0.032	0.945	2.768
{OCM-1, WY-1}	{AF-1}	0.054	0.945	2.766
{A-5, OCM-1}	{AF-1}	0.033	0.944	2.764
{CS-1, OCM-1, WY-1}	{AF-1}	0.033	0.944	2.766
{OCM-12}	{AF-1}	0.029	0.916	2.684
{OCM-13}	{AF-1}	0.027	0.886	2.594
{MTP-4, OCM-3}	{AF-1}	0.026	0.857	2.51
{OCM-11}	{AF-2}	0.027	0.802	4.906

Table 7은 637개의 규칙 중에서, 지지도 0.025 이상, 신뢰도 0.8 이상, 향상도 2.5 이상을 동시에 만족하는 총 22개의 특징적인 규칙을 나타낸다. 여기에서 지지도가 가장 높은 규칙은 “사다리(OCM-1)”에서의 “떨어짐(AF-1)”이지만, 신뢰도와 향상도가 동시에 가장 높은 규칙은 “13시~16시(AT-3)”에 “바닥·통로(OCM-2)”에서 “넘어짐(AF-2)”임을 알 수 있다. 떨어짐에 대하여 신뢰도와 향상도가 가장 높은 규칙은 “공사규모 3억 미만(CS-1)”인 현장에 “사다리(OCM-1)”에서 “떨어짐”으로 나타났다. 그 외의 규칙들을 살펴보면, “공사규모 3억 미만(CS-1)”, “연령 50대”, “기타공사(CT-7)”와 “건축공사(CT-1)”, “근속기간 1개월 미만”은 떨어짐과 넘어짐 재해에 공통적으로 관계된 요소임을 알 수 있다. 따라서, 공사규모 3억 미만의 건축공사나 기타공사 현장에서는 근속기간 1개월 미만 50대 근로자의 사다리에서의 떨어짐이나 바닥·통로에서의 넘어짐에 대한 안전관리를 철저히 할 필요가 있다. 한편, “요양기간 3~6개월(MTP-4)”은 떨어짐 재해에서, “요양기간 1~3개월(MTP-5)”은 넘어짐 재해에서만 나타나는데, 이는 떨어짐이 넘어짐에 비해 더 심한 부상을 입을 가능성이 높다는 것을 시사한다. 또한 넘어짐 재해에서는 “건설기능인력(O-1)”과 “단순노무자(O-2)”가 주요한 직종으로 나타나지만, 떨어짐은 직종에 대한 항목이 나타나지 않았다는 점에서 떨어짐은 직종에 관계없이(즉, 모든 직종에서) 발생할 가능성이 높다고 볼 수 있다.

#### 4.3 Pareto 분석과의 비교

Table 8은 구해진 637개의 규칙에 포함된 항목들의 종류를 나타낸다. 즉, Table 8에 나타난 항목들은 Table 2에서 제시한 항목들 중에서도 건설공사의 비사망 재해에 있어서 연관 규칙적 패턴을 보이는 특징적인 요소에 해당한다고 볼 수 있다. 이를 앞의 Pareto 분석 결과와 비교해보면, 대부분의 항목들이 Pareto 분석에서 상위 빈도를 차지하는 항목들에 해당함을 알 수 있다. 그러나, Pareto 분석에서 상위 빈도 항목에 해당하지만 도출된 규칙에 포함되지 않는 항목이나, 반대로 하위 빈도의 항목에 해당하지만 도출된 규칙에 포함된 경우도 일부 존재하고 있음을 알 수 있다. 이러한 결과는 Pareto 분석과 같이 단순 빈도 위주의 통계 분석에서는 발생 가능성이 낮은 것으로 평가되는 요소일지라도, 이러한 요소를 다른 요소들과의 상호 관계를 기준으로 분석하면 다른 요소들에 비해 더 높은 발생 가능성을 보일 수 있다는 것을 나타내며, 이는 효과적인 안전관리 대책 수립을 위한 재해 분석에 있어 ARM 분석의 유용성을 입증한다.

Table 8. Characteristic factors included in ARM results

Variables	Factors included in	
	If (X)	Then (Y)
Occupations	O-1, O-2	O-1
Accident times	AT-1, AT-2, AT-3, AT-4	-
Medical treatment periods	MTP-3, MTP-4, MTP-5	-
Ages	A-5, A-6	-
Facilities	F-1, F-2, F-3, F-4, F-5	F-1
Construction types	CT-1, CT-7	CT-1, CT-7
Accident forms	AF-1, AF-2, AF-3, AF-4	AF-1, AF-2, AF-3, AF-4
Original cause materials	OCM-1, OCM-2, OCM-3, OCM-4, OCM-5, OCM-6, OCM-7, OCM-8, OCM-10, OCM-11, OCM-12, OCM-13, OCM-15, OCM-17	OCM-2, OCM-4
Working years	WY-1	WY-1
Process rates	PR-1, PR-2, PR-3, PR-4	-
Construction scales	CS-1, CS-2	CS-1, CS-2

한편, Table 8에서 조건절(X)에 나타난 항목들은 Table 2의 각 변수의 세부 항목들이 다수 포함되어 있으나, 결과절(Y)에 나타난 항목들에는 일부 변수의 항목이 포함되어 있지 않을 뿐만 아니라, 조건절의 항목 중에서도 특정 항목만이 포함된 것을 알 수 있다. 이러한 결과는 ARM이 단순한 상관관계보다는 신뢰도 지표에 의해 어느 정도 인과성이 있는 관계를 탐색하기 때문으로 사료된다. 예컨대, 조건절에는 기인물(OCM)이 포함되고, 결과절에는 재해형태(AF)가 포함된 규칙이 전체 637개의 규칙 중에서 322개(50.05%)를 차지한다는 점이 이를 입증한다. ARM 분석 결과에서 나타나지 않은 항목들은, 이들을 포함하는 규칙이 본 연구에서의 분석 조건 (지지도 0.015 이상, 신뢰도 0.5 이상)을 만족하지 않기 때문에 나타나지 않은 것으로 판단된다. 예를 들어 Table 5에 제시된 재해 유형 중에서, AF-8 이하의 유형들의 Pareto 빈도는 최소지지도(0.015)보다 낮은 값을 가지므로, 이러한 경우(즉, 최소지지도 이하의 Pareto 빈도를 갖는 항목이 포함되는 규칙)은 도출 규칙에 포함되지 않았다. 또한, AF-5나 AF-6와 같이 Pareto 빈도는 최소지지도보다 높지만, 이를 포함하는 규칙이 신뢰도를 만족하지 못할 경우에도 규칙으로 나타나지 않았다. 예를 들어, 본 연구에서 AF-5와 A-5가 동시에 나타나는 건수는 전체 59,202건 중에서 2,136건이며, 따라서 이 경우의 지지도는 0.036으로 최소지지도 이상을 갖지만, 신뢰도는 {AF-5⇒A-5}일 때 (0.036/0.088 = 0.409), {A-5⇒AF-5}일 때 (0.036/0.4256 = 0.0846)으로 모두 최소 신뢰도인 0.5 이



하의 값에 해당하므로, 이러한 경우 분석된 규칙에는 포함되지 않는다. 따라서, Table 8에 나타나지 않은 항목들은, 본 연구에서의 ARM 분석 조건 내에서는 다른 변수들의 항목들과 어떠한 규칙성을 갖지 않고 무작위(Random)하게 발생할 가능성이 높은 항목이라고 볼 수 있다.

### 5. 결론

본 연구에서는 건설공사의 사고성 산업재해의 대부분을 차지하는 비사망재해에 영향을 미치는 여러 잠재 요인들 간의 연관 특성을 알아보려고 하였다. 이를 위해 59,202건의 건설업 산업재해 데이터에 대해 최소지지도 0.015, 최소신뢰도 0.5로 설정하여 ARM 분석을 실시하였고, 총 637개의 연관규칙을 도출하였다. 도출된 규칙은 Pareto 분석 결과와 함께 비교 분석하였다. 이를 통해 얻은 결론은 다음과 같다.

(1) 지지도 기준 상위 규칙들은 Pareto 분석에서 상위의 빈도를 차지하는 항목들이 단순한 형태로 제시되는 규칙들이 상위 규칙으로 많이 나타나지만, 항상도나 신뢰도 기준의 상위 규칙들은 Pareto 분석에서 상위에 있는 항목들 뿐만 아니라 하위에 해당하는 항목들도 포함하여 좀 더 복합적인 형태의 규칙들이 상위 규칙에 많이 나타난다.

(2) 건설공사의 비사망재해에 대해 지지도, 신뢰도, 항상도가 모두 높은 규칙은 사다리에서 떨어짐과 바닥·통로에서 넘어짐과 관련한 규칙들이며, “공사규모 3억 원 미만”, “연령 50대”, “건축공사”와 “기타공사”, “근속기간 1개월 미만”은 떨어짐과 넘어짐 재해에 공통적인 조건에 해당한다. 또한, 넘어짐 재해는 건설기능인력과 단순노무자 직종에서 발생할 가능성이 높지만, 떨어짐 재해는 직종과 관계없이 발생할 가능성이 높다.

(3) Pareto 분석은 다른 항목들과의 결합 발생 가능성을 정량적으로 파악할 수 없으나, ARM은 연관성이 있는 다른 항목들이 무엇인지 뿐만 아니라 연관성이 있는 항목들과의 결합 발생 가능성도 정량적으로 파악할 수 있다. 따라서, 안전관리의 우선 순위를 결정하기 위한 재해분석시에는 ARM과 같은 복합 분석을 수행하는 것이 필요하다. 예를 들어, {사다리(OCM-1)}⇒{떨어짐(AF-1)} 규칙 보다, {공사규모 3억 미만(CS-1), OCM-1}⇒{AF-1} 규칙이 더 높은 발생 가능성을 갖는 것으로 나타났다. 이는 3억 미만 소규모 현장에서는 사다리가 떨어짐 재해의 다른 기인물보다 특히 더 주요한 기인물임을 나타내며, 공사규모 3억 미만 소규모

현장은 사다리에서의 떨어짐에 대한 안전관리를 더욱 철저히 할 필요가 있다는 점을 시사한다.

(4) Pareto 분석에서 발생 빈도가 낮은 요소일지라도, 이러한 요소를 다른 요소들과의 상호 관계를 기준으로 분석하면 다른 요소들에 비해 더 높은 발생 가능성을 보일 수 있다. 예를 들어, Table 4에서 OCM-10은 OCM-9보다 건설공사 비사망재해에 대한 Pareto 빈도가 낮은 기인물 항목이나, Table 8에 나타난 바와 같이 OCM-10은 다른 항목들과의 연관규칙이 나타난 반면에 OCM-9에 대해서는 다른 항목들과의 연관규칙이 분석 결과에서 나타나지 않았다. 이는 건설공사 비사망재해에 있어서 OCM-10의 발생 빈도는 OCM-9보다 낮지만, OCM-10이 다른 요소들과 결합될 경우에는 OCM-9보다 발생 가능성이 높다는 것을 시사한다.

### References

- 1) Ministry of Employment and Labor (MOEL), ‘2020 Industrial Accident Status Analysis’, 2021.
- 2) H. S. Lee, H. S. Kim, M. S. Park, K. P. Lee and S. B. Lee, “Construction Risk Assessment Methodology Using Site Risk Influence Factors”, Korean Journal of Construction Engineering and Management, Vol. 10, No. 6, pp. 117-126, 2009.
- 3) J. Hinze, ‘Construction Safety’, Prentice Hall, NJ, USA, 2006.
- 4) J.M. Hwang and S.W. Shin, “Correlational Structure Modelling for Fall Accident Risk Factors of Portable Ladders Using Co-occurrence Keyword Networks”, J. Korean Soc. Saf., Vol. 36, No. 3, pp. 50-59, 2021.
- 5) K.Y. Son and H.G. Ryu, “Association Rules Analysis of Safe Accidents Caused by Falling Objects”, Journal of Korean Institute of Building Construction, Vol.19, No. 4, pp. 341-350, 2019.
- 6) G.H. Lee, C.S. Lee, C.W. Koo and T.W. Kim, “Identification of Combinatorial Factors Affecting Fatal Accidents in Small Construction Sites: Association Rule Analysis”, Korean Journal of Construction Engineering and Management, Vol. 21, No. 4, pp. 90-99, 2020.
- 7) J.S. Lim, S.U. Han, Y.C. Kang, S.H. Kang, “Affinity Analysis Between Factors of Fatal Occupational Accidents in Construction Using Data Mining Techniques”, Korean Journal of Construction Engineering and Management, Vol. 22, No. 5, pp. 29-38, 2021.
- 8) B. Kim, S. Chang and Y. Suh, “Text Analytics for

- Classifying Types of Accident Occurrence Using Accident Report Documents”, J. Korean Soc. Saf., Vol. 33, No. 3, pp. 58-64, 2018.
- 9) S. Kang S. Chang, J. Lee and Y. Suh, “Structuring Risk Factors of Industrial Incidents Using Natural Language Process”, J. Korean Soc. Saf., Vol. 36, No. 1, pp. 56-63, 2021.
- 10) J. Weng, J. Z. Zhu, X. Yan and Z. Liu, “Investigation of Work Zone Crash Casualty Patterns Using Association Rules”, Accident Analysis and Prevention, Vol. 92, pp. 43-52, 2016.
- 11) C. W. Cheng, C. C. Lin and S. S. Leu, “Use of Association Rules to Explore Cause-Effect Relationships in Occupational Accidents in the Taiwan Construction Industry”, Safety Science, Vol. 48, Issue 4, pp. 436-444, 2010.
- 12) C. Zhang and S. Zhang, ‘Association Rule Mining: Models and Algorithms’, Springer, N.Y., USA, 2002.
- 13) Korea Occupational Safety and Health Agency(KOSHA), ‘Evidence Based Accident Analysis’, KOSHA Report (2017-Construction-741), 2017/
- 14) J. H. Won, H. D. Lee, K. J. Song, and S. H. Jung, “Analysis of Construction Accident Characteristics during Past 10 Years”, Proceedings of Korean Society of Disaster and Security Conference, pp. 140-145, 2016.