A MULTIMODAL APPROACH TO IMPROVE FIRE SAFETY ON CONSTRUCTION SITES

By

JAEHONG KIM

Bachelor of Architectural Engineering
Kyungpook National University
Daegu, South Korea
2010

Master of Architectural and Civil Engineering
Kyungpook National University
Daegu, South Korea
2012

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2022

A MULTIMODAL APPROACH TO IMPROVE FIRE SAFETY ON CONSTRUCTION SITES

Dissertation Approved:

Yongwei Shan

Dissertation Adviser

Jong Hoon Kim

Mohamed Soliman

Haejun Park

Name: JAEHONG KIM

Date of Degree: DECEMBER, 2022

Title of Study: A MULTIMODAL APPROACH TO IMPROVE FIRE SAFETY ON CONSTRUCTION SITES

Major Field: CIVIL ENGINEERING

Abstract: Fire risk on construction sites is a threat to property and construction workers' safety. Fire safety on construction sites has been rarely studied because fire accidents have a lower occurrence compared to Construction's "Fatal Four". Despite the lower occurrence, construction fire accidents tend to have a larger severity of impact. To minimize the loss of human life and property, effective fire safety plans should be established on the construction site. This dissertation is composed of three studies. The first study aims at using news media and big data analysis to identify factors related to fire accidents on construction sites. The authors identified the level of media exposure for various keywords related to construction accidents and analyzed the similarities between keywords. The results show that the level of media exposure for fire accidents on construction sites is much higher than for fall accidents, which suggests that fire accidents may have a greater impact on the surroundings than other accidents. The second research was conducted to identify important influencing factors on fire safety on construction sites as well as their relationships through a questionnaire administered to 1,492 fire officers. The authors identified the interrelationships between factors through path analysis and structural equation modeling. The paths of factors showed the highest direct relation between building and fire factors and that the human factor played a mediator role between human and fire factors. The last study aims to provide a BIM-based simulation framework to simulate fire and evacuation performance on construction sites. Computer simulation can be a cost-effective approach for calculating evacuation performance and fire growth on a construction site. To propose a framework, the author analyzes how to apply BIM to fire safety simulations and classifies BIM-based input data required for simulation. This study contributes to the body of knowledge by exploring factors related to fire safety on construction sites and their interrelationships. This study also contributes by effectively integrating fire safety into construction sites and providing practitioners with fire accident analysis data. This study can play an important role in establishing the regulations necessary to improve safety on construction sites.

TABLE OF CONTENTS

Chapter Page

Chapter Page

Chapter Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION AND BACKGROUND

PROBLEM STATEMENT

Due to the continuous development of construction technology, various types of construction projects are in progress. As construction projects diversify, activities on the construction site become more complicated. The complexity of these construction projects can lead to increased risk and accidents on the site. According to The Bureau of Labor Statistics (BLS), 5,250 fatal work injuries and 1,008 worker deaths on construction sites were recorded in 2018, this is a 2% increase from 2017. Accidents on construction sites cause significant financial damages as well as personal injuries. In the National Fire Protection Association (NFPA) report, the frequency of fires related to construction among all building fires is reported to be about 1%, but direct property damage is reported to be about 2%. This shows that accidents on construction sites cause more financial damage than frequency. In addition, accidents on the construction site can lead to an extension of the construction period, which greatly affects the management of the construction project.

Fire accidents are an important type that can affect the safety of a construction site, but most fire safety studies focus on the building in use. Research is often conducted on evaluation programs and surveys to enhance the fire safety of buildings in use. In addition, the software is being developed to evaluate and manage building fire safety levels. However, these studies have limitations that reflect the knowledge and opinions of fire accidents at construction sites in some occupational groups, such as construction workers. Many studies use surveys from construction site managers and workers to analyze the risks on construction sites. They are experts in construction-related areas, but there is a possibility that they are not experts in accident and risk management such as a fire. In order to overcome these limitations, research based on the experience and knowledge of various experts related to accidents on the construction site must be conducted. In particular, in the case of fire accidents on construction sites, the experience and knowledge of firefighters are essential for related research. Firefighters are the most experienced occupations in fire accidents, and they have accumulated fire-related knowledge through numerous education and training. However, few studies have used firefighters' perceptions or experiences in research on fire safety on construction sites.

Many construction site accident-related studies focus on types with a high frequency of accidents in the past. However, the limitations are clear in studies that analyze only the frequency of accidents. Although the frequency of fire accidents on construction sites is low, it can have a greater impact on construction sites than other types of accidents. This is because fire accidents have the character of spreading to the surroundings, such as explosions and spreads. In order to accurately analyze the damage of each type of accident on the construction site, it is necessary to consider the impact of each type of accident on the construction project and the surrounding environment. The impact on construction projects and the surrounding environment includes various external conditions such as damage to surrounding buildings and the extension of construction periods, making it difficult to quantify these impacts. For this reason, most construction-related studies focus only on the frequency of accidents. In order to quantify the impact of each accident, it is necessary to utilize various methodologies used in other fields for research. The frequency of media exposure can be used as an indirect material to quantify the impact of accidents on construction sites. The media is characterized by dealing with accidents that have a greater social impact than minor injuries to construction sites. There is a high possibility that the media will be exposed to accidents that can seriously affect the surrounding environment and construction projects. In the related research, these media characteristics can be used to analyze the impact of each accident type on the construction site.

In addition, various new methodologies such as web crawling, deep learning, and word embedding can be used to analyze factors that can affect fires on construction sites. Studies related to accidents on construction sites are limited to studies that analyze the causes of past accidents and analyze the causes, or measure risks in the field through surveys of some construction project stakeholders. This trend in the field of construction safety contrasts with other fields that have been actively used in various data analysis methodologies. If researchers in the field of construction safety accept new changes and attempt to analyze through various methodologies, they will be able to present meaningful results that can improve safety on construction sites. Also, It will also help researchers to understand a lot of relevant readers when presenting and visualizing the analyzed data more intuitively than providing a simple table.

Finally, BIM data is used to maintain the quality of the design on the construction project. New regulations have been established that require the use of BIM data in administration parts such as governments to verify building permits and regulations. However, the use of BIM data on construction sites tends to focus only on the quality of the design. Since BIM data contains a large amount of information such as schedules, materials, and topologies in construction projects, it can be used to improve fire safety in construction sites. Currently, a study that analyzes accidents on a construction site using BIM data has a limit to check only the design plan and rules, such as finding errors in the design and eliminating the risk of the site. One of the reasons for this narrow research trend is that designers are not provided with clear data definitions and guidelines for safety and risk analysis of construction sites when generating BIM data. To effectively use BIM data to enhance construction site fire safety, it is very important to define BIM data structure for site fire safety and provide it to designers. In order to compensate for this limitation, this study presented a new analysis method using survey and data analysis.

RESEARCH BACKGROUND

Case study of fire accidents on construction sites

Fires and explosions can occur at construction sites by unfinished fire protection systems, sparks from the welding operation, leaking gases, incomplete electrical systems, arson, and so on. In construction sites, the fire evacuation system is not equipped properly, and it is difficult to detect the fire early and to have first-aid firefighting. Moreover, many loads and heavy vehicles into and out of the construction sites can make it difficult for fire trucks to gain access. For these reasons, if a fire occurs during construction, the loss can be larger than the fire which breaks out in the completed building. Therefore, it is necessary to ensure the fire safety of construction sites to minimize damage or loss. The chapter below describes some examples of fire accidents at construction sites in the U.S. and South Korea.

According to the Occupational Health & Safety Administration (OSHA), 200 construction workers are killed and over 5,000 construction workers are injured by fires and explosions across all industries each year in the U.S. Among them, there are on average 36 fire and explosion deaths per year in the construction industry and over the last 11 years, OSHA recorded 361 fire or explosion deaths in the construction industry. Most of the fire accidents at construction sites happen because the flammable materials of the site were ignited by welding and cutting sparks. For instance, in January 2018, a fire blazed on an apartment complex under construction near San Diego state university. The fire was ignited by the use of stucco paper. Two of the three buildings under construction were burned. The roofer didn't know that the stucco paper was flammable. There is another example of a fire caused by workers without proper safety training. The fire broke out at demolition works on Deutsche Bank which is a 39-story high-rise in downtown Manhattan, New York in August 2007. The main cause of this accident was the workers' cigarettes. Two firefighters were killed, and 50 others were injured in the accident. Firefighting activities on construction sites can be delayed due to incomplete fire prevention plans and building materials. An example is a fire accident on the construction site of a development project in San Francisco in March 2014. Over 150 firefighters were brought in to control the fire but 172 units under construction were burned. This site had no sprinkler system therefore it caused much more damage. Fire in construction sites requires more effort from firefighters due to the flammable building materials. For instance, the fire occurred at the apartment building in College Park which was under construction for more than 270 housing units near the University of Maryland in April 2017. The fire wounded 4 people and the damage cost more than $39 million. It also required more than 200 firefighters and the largest fire suppression effort because the building had so much wood material.

Figure 1-1 Examples of fire accident on construction site in the U.S.

According to the Ministry of Public Safety and Security (MPSS) of South Korea, one of the main causes of fire on construction sites is welding sparks like in the U.S. About 1,075 fire accidents on construction sites break out per year and 2.9 fires per day are caused by welding. For instance, there was a fire in a 30-story apartment building in Seoul. The fire broke out because welding sparks were set on the urethane foam. The firefighters succeeded in extinguishing the fire in 24 minutes, but one person died because of the toxic gas from urethane foam. Unlike other developed countries, Korea's construction sites widely use a lot of organic insulation such as urethane foam instead of fire-retardant insulation due to its low cost. However, the toxic gas from this kind of insulator can lead to massive damage if there is a fire. The leading cause of death is caused by toxic gas suffocation, rather than burns in South Korea. In December 2008, the new construction site of the GS warehouse was burnt out due to the welding sparks. Since the main material in the warehouse was a sandwich panel capable of being vulnerable to fire and generating toxic gases, seven people died and two were injured. Moreover, In May 2014, a large public transportation network center caused a fire from welding, resulting in eight deaths and 58 injuries. Also, Incomplete structure and insufficient site information can also lead to a lot of damage from fire as well as a toxic gas. Since many workers are daily employees in South Korea, they are not familiar with the structure and safety guidelines of the construction site. Thus, it can also lead to more serious damage if there is a fire. For example, when a fire broke out in a new construction site of cold storage, Icheon in 2008, 40 people died, and nine people were injured. Since the fire broke out in confined spaces and there was only one exit, this accident was deadly.

Figure 1-2 Examples of fire accident on construction site in South Korea

Table 1-1 shows an overview of fire accidents at construction sites that are issued in this chapter. Through this case study, it seems that there are several efforts needed in order to ensure safety from the fires. First, the construction manager should provide fire safety training for workers properly. Not only safety training but also continued modification of construction safety regulations which may cause fires is also needed. The modification should consider the development of relevant technology and the characteristics of construction sites. Second, the construction manager has to seek a way for extinguishing the fire in the beginning. The initial suppression of a fire is important to extinguish the fire and it can minimize damage caused by fires. For that, the site manager should monitor carefully and continuously and should perform checks on the firefighting equipment that is used in the construction sites. Finally, the owner and contractor of the construction projects need to consider using non-combustible materials as well as construction costs. These efforts are able to improve safety for workers and reduce the damage caused by fire accidents on construction sites.

Table 1-1 Examples of fire accident at construction site

The U.S.		
Date	**Place**	**Details, caused and damage of accident**
August, 2007	Manhattan, New York	- A cigarette of worker was set on the toxic debris floor - Two firefighters were killed and 50 were injured
March, 2014	Mission Bay, San Francisco	- Caused by welding spark - 172 units under construction were burned - One fire fighter was injured - Over 150 firefighters were dispatched
March, 2014	Houston, Texas	- Caused by welding spark - The estimated damage was over $ 50 million - More than 200 firefighters and 80 units were dispatched
April, 2017	University of Maryland, Maryland	- More than $39 million worth was damaged - Four people including a firefighter were injured - Over 200 firefighters were dispatched
July, 2017	Waltham, Massachusetts	- Caused by arson - The estimated damage was about $110 million

		- Two firefighters had minor injuries - More than 100 people were evacuated
January, 2018	San Diego State University, San Diego	- Caused by torch sparks - Two of three buildings under construction were burned - One worker had minor injury
June, 2017	Dorchester, Boston	- Caused by the exhaust pipe installed improperly - The estimated damage was more than $45 million - One worker was treated for chest pain.
South Korea		
Date	**Place**	**Details, caused and damage of accident**
March, 2007	Seoul	- Caused by welding spark - One people died, 60 workers were injured - The estimated damage was 230 million won
January, 2008	Icheon	- Caused by welding spark - 40 people died, nine people were injured - The estimated damage was 600 million won.
December, 2008	Icheon	- Caused by welding spark - Seven people died, two people were injured - The estimated damage was more than 53 billion won
August, 2012	Seoul	- A short circuit caused a fire - Four people died, 25 people were injured
May, 2014	Goyang	- Caused by welding spark - Eight people died and 58 people were injured - 95 units, 334 firefighters were dispatched to the scene
June, 2016	Gyeonggi	- A gas leak may have caused the blast - Four people died and 10 people were injured
September, 2016	Gimpo	- Caused by welding spark - Four people died and two people were injured.
December, 2017	Gwanggyo	- Caused by welding spark - One people died, 14 people were injured - Six fire helicopters, 59 units, and 138 firefighters of adjacent fire station were dispatched to the scene

Statistics of the causes of fire accidents on construction sites

From 2010 to 2014, the National Fire Protection Association (NFPA) investigated the number of average fires, the causes of fire, and the losses caused by a fire on construction sites each year. According to NFPA, on average 485,700 fires break out per year in the U.S. On average 2,614 people die, 14,209 people are injured, and $8,439 million is directed to property damage by fire per year. Of that, the number of average fires per year on construction sites is 8,440. On average, the fire caused 13 dead workers, 132 injured workers, and $310 million of damage to construction sites. More specifically, construction sites can divide into three types; under construction, under renovation, and under demolition. Firstly, there was an average of 3,750 fires, 5 deaths, 51 injuries, and $172 million of damage that occurred every year in the buildings under construction. Secondly, on average 2,560 fires broke out in buildings under renovation, and these fires killed four people, wounded 65 people, and generated $108 million of damage per year. Finally, in the case of the building under demolition, four people died, 16 people were injured, and $30 million in damage occurred caused by 2,130 fires. Overall, fires occur most frequently in the building under construction and the greatest direct property damage from the fires is also in the structure under construction stage. A possible explanation for this might be that many activities take place simultaneously in the building under construction and those activities can increase the danger of fires. In addition, there is a high probability of loading combustible materials for construction in the building compared to renovation and demolition. Table 1-2 below shows the frequency and damage of fires on construction sites by year.

Table 1-2 Construction site fire in 2010-2014 annual averages

Status	Fires	Civilian Deaths	Civilian Injuries	Direct Property Damage (in Millions)
Under construction	3,750	5	51	$172
Undergoing major renovation	2,560	4	65	$108

Being demolished	2,130	4	16	$30

According to the NFPA investigation, the cooking equipment (27%) is the biggest cause of fire in the building under construction. In the case of the fire in the site under renovation, the leading cause is heating equipment (15%). Lastly, there were a lot of intentional fires (42%) in the structure under demolition. The main causes are different depending on the type of construction site. Thus, it can be suggested that people need to update rules regularly for the construction site fire safety based on these investigated leading causes. However, the current fire statistics from the NFPA did not investigate with a classification of characteristics of the construction site. Therefore, more investigation considering the characteristics of construction sites such as season, size of the site, and so on is needed to form a better analysis. Figures 1-3, 1-4, and 1-5 provide the main cause of the fire by the construction stage.

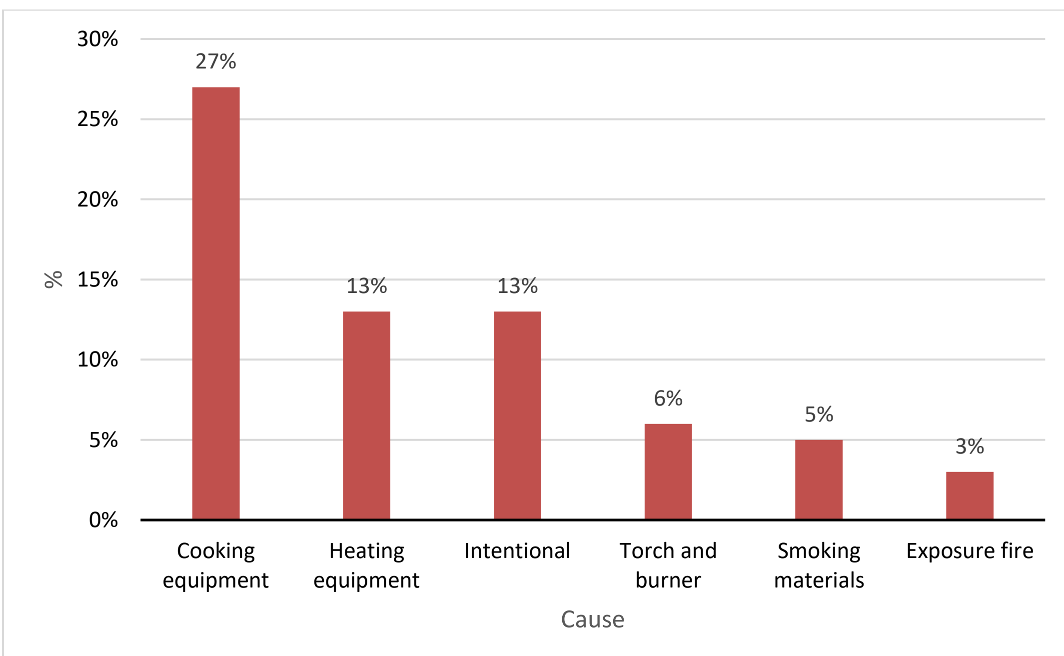

Figure 1-3 Causes of construction site fire accidents (under construction)

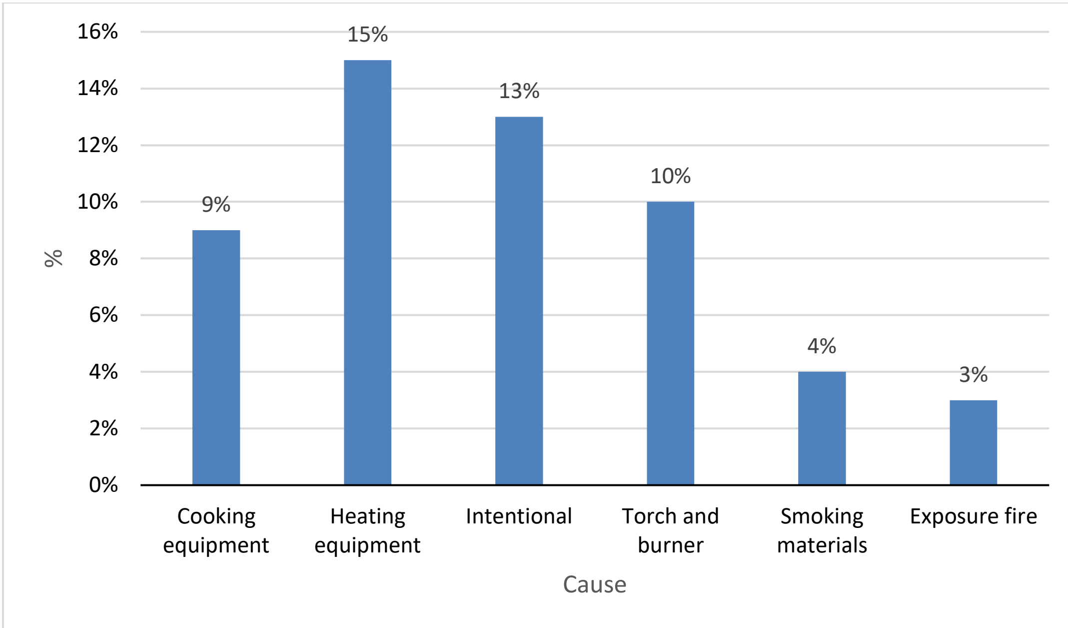

Figure 1-4 Causes of construction site fire accidents (major renovation)

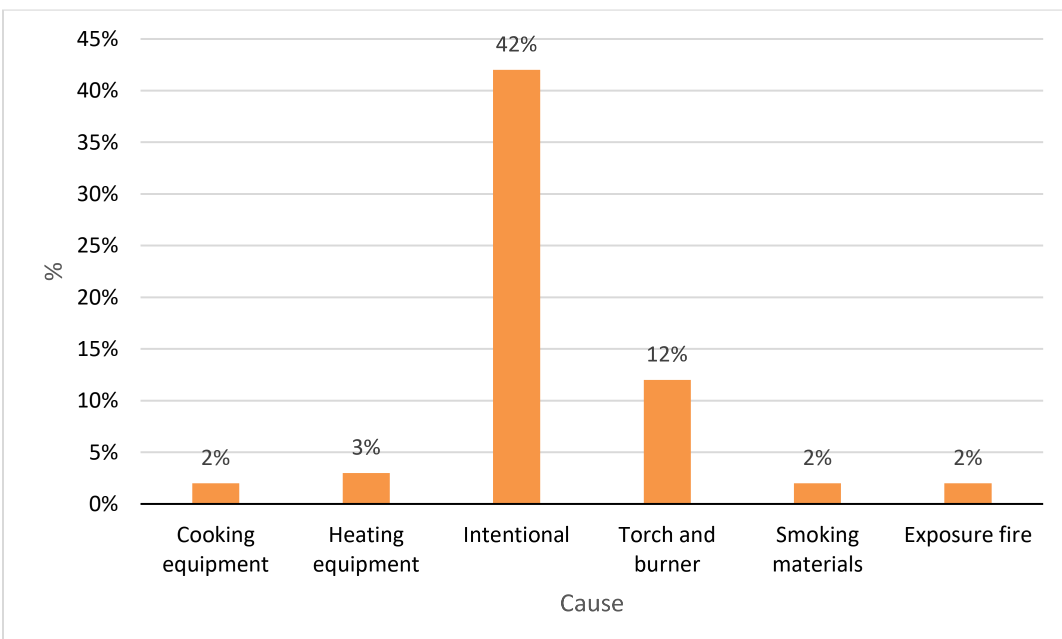

Figure 1-5 Causes of construction site fire accidents (being demolished)

DISSERTATION STRUCTURE

This dissertation is composed of three journal papers and includes four chapters. The first chapter includes the problem statement, research background, dissertation format, and research objectives. Each of the next three chapters provides an independent article. Chapters 2, 3, and 4 follow in the ASCE journal paper format. Therefore, each chapter has its own abstract, introduction, background, methodology, results, discussion, and conclusion. In Chapter 2, a questionnaire was conducted to investigate the perception of firefighters on a construction site fire. The questionnaire was conducted for 1,492 firefighters working in South Korea. The author found out the interrelationships between fire-related factors through frequency analysis and path analysis. The results provide the relationship between the factors of the fire accident from the firefighter's aspect, which is important for construction site workers to identify the fire hazard. Chapter 3 provides the impact of fire accidents on construction sites was investigated using the frequency of media exposure. The media is always interested in relatively significant events, so there is very little chance of documenting minor accidents on the construction site. On the other hand, larger-scale accidents that have a huge impact on the surroundings are likely to be reported in news articles. By utilizing these characteristics of the media, the impact and relationship of fire accidents on construction sites were identified. Using the collected data, the author analyzed the frequency of factors related to construction site accidents. In addition, similarity and relationships between related factors were analyzed through word embedding and network analysis. To intuitively visualize each word that has a high relationship between fire accident and fall accident, the Uniform Manifold Approximation and Projection (UMAP) method is applied. Finally, in Chapter 4, the authors provided a BIM-based simulation framework to analyze fire and evacuation performance on construction sites. To provide a framework, the author analyzes how to apply BIM in fire safety simulations and classifies BIM-based input data required for simulation. Also, the IFC data structure was analyzed to provide guidelines for applying BIM to improve fire safety on construction sites. This study is a new approach to effectively applying BIM to construction site safety. In addition, by providing BIM guidelines related to fire safety on construction sites, BIM data can be used more effectively to improve site safety.

RESEARCH OBJECTIVES

The conceptual overview of this research objective is presented in the following figure 1-6. The primary objective of this research is to develop new approaches to improve fire safety on construction sites by applying various data mining concepts and techniques such as path analysis, web crawling, word embedding, deep learning UMAP, and BIM data analysis. This main purpose can be divided into three secondary objectives, as follows:

1. Analysis of Fire Accident Factors on Construction Sites Using Web Crawling and Deep Learning Approach
2. Examine Factors Influencing Fire Safety on Building Construction Sites: A Fire Officer's Perspective
3. Developing Data Requirements of BIM for Fire Safety on Construction Sites

The contribution of this study is to effectively integrate fire safety into the construction site and provide practitioners with data related to fire accident prevention. A new approach was presented using firefighter's perception, web crawling, word embedding, and BIM data, which are methodologies that have not been actively utilized in research related to fire safety on construction sites. Also, this study can play an important role in establishing the regulations necessary to increase safety in construction sites.

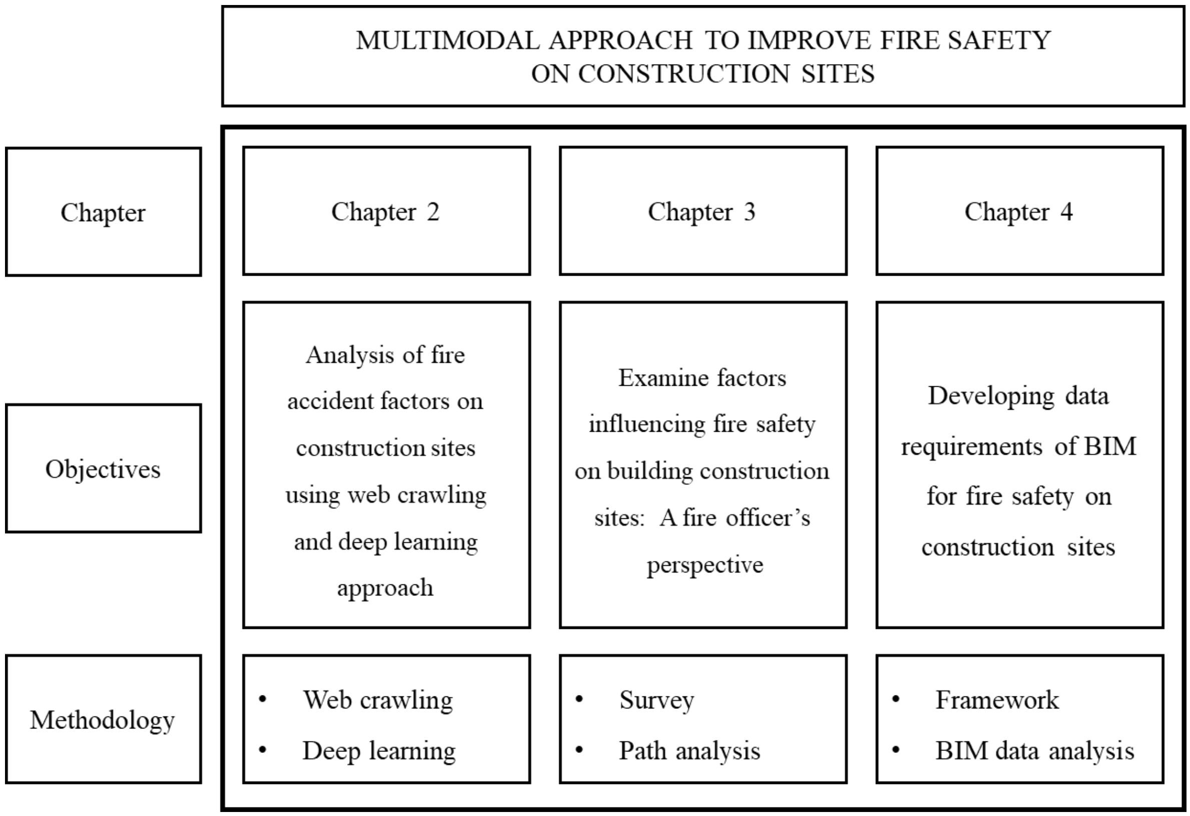

Figure 1-6 Conceptual overview of dissertation

CHAPTER II

ANALYSIS OF FIRE ACCIDENT FACTORS ON CONSTRUCTION SITES USING WEB CRAWLING AND DEEP LEARNING APPROACH

ABSTRACT

Fire safety on construction sites has been rarely studied because fire accidents have a lower occurrence compared to Construction's "Fatal Four". Despite the lower occurrence, construction fire accidents tend to have a larger severity of impact. This study aims at using news media data and big data analysis techniques to identify patterns and factors related to fire accidents on construction sites. News reports on various construction accidents covered by news media were first collected through web-crawling. Then, the authors identified the level of media exposure for various keywords related to construction accidents and analyzed the similarities between keywords. The results show that the level of media exposure for fire accidents on construction sites is much higher than for fall accidents, which suggests that fire accidents may have a greater impact on the surroundings than other accidents. It was found that the main causes of fire accidents on construction sites are violations of fire safety regulations and the absence of inspections, which could be sufficiently prevented. This study contributes to the body of knowledge by exploring factors related to fire safety on construction sites and their interrelationships as well as providing evidence that the fire type should be emphasized in safety-related regulations and codes on construction sites.

INTRODUCTION

Fire accidents are greatly affected by external environments, such as weather and surrounding buildings or hazards, making it difficult to control and prevent (Caton et al. 2017). This is especially dangerous on building construction sites because fire safety equipment such as sprinklers and fire alarms may not be completed depending on the progress of the construction (Hamid et al. 2003). Most construction accident-related research focuses on Construction's "Fatal Four" as they occur more frequently. The Fatal Four Hazards provided by the Occupational Safety & Health Administration (OSHA) consist of falls, electrical exposure, struck-by, and caught-in/between. Since the Fatal Four is based on the frequency of accidents on construction sites, fire accidents are excluded. In addition, it has a limitation in not considering various impacts such as the secondary spread of accidents. However, fall accidents, for example, on construction sites are less likely to lead to secondary accidents. Therefore, it is necessary to examine the impact of fire accidents on construction sites and their surroundings in spite of their low occurrence. As a result, new fire safety regulations or rules for construction sites might be proposed.

The news media often covers accidents that have a great social impact on construction sites. Fatal accidents, such as major fire accidents, that affect the surroundings are more likely to be exposed to the media. Therefore, the lessons from the articles provided by media can be learned. By measuring media exposure by accident type, it can be a new attempt to measure the impact of each type of accident. Moreover, the articles provided by the media are organized in a similar format, which is efficient for researchers to use the data. To explore the level of impact of fire accidents on construction sites, this study collected articles on construction site accidents reported in the New York Times over the past 20 years. The web-crawling method was used for efficient and accurate data collection. In addition, to analyze and visualize the relationship between factors related to fire safety on construction sites, word embedding, network analysis, and Uniform Manifold Approximation and Projection method were applied. This study contributes to the body of knowledge by exploring the factors of construction site fire and providing relationships between factors through the collection of media big data on fire accidents on construction sites. In addition, this study promotes the active inclusion of fire types in the aspect of safety on construction sites based on the results of the study.

BACKGROUND AND LITERATURE REVIEW

The construction industry always considers safety, but the fatality rate on construction sites remains high (Abdullah and Wern 2011). According to the Occupational Safety & Health Administration (OSHA), 20.5% of fatal workplace accidents occur on construction sites (Hosseinian and Torghabeh 2012). In particular, a fire accident on a construction site can lead to secondary accidents such as fire spread and explosion, which can greatly increase the project cost and duration. According to the National Fire Protection Association, annual direct damage to US construction sites due to fire accidents on construction sites is estimated at about 172 million dollars (Campbell 2017). When the direct damage to the construction site related to renovation and demolition is combined, it amounts to 300 million dollars per year. To reduce construction site fire accidents, various studies have been conducted to analyze accidents on construction sites, but there are limitations. Many studies related to fire accidents on construction sites are limited to analyzing the causes of accidents individually (Ahn et al. 2018; Ali et al. 2010). In order to increase fire safety on construction sites, it is necessary to collect big data related to fire accidents on construction sites and analyze the factors of fire using various approaches.

Construction site fire safety

There are various threats on construction sites that can lead to fire accidents, such as flammable materials and dangerous construction activities (Maluk et al. 2017). Fire accidents on construction sites are likely to cause secondary accidents such as collapse, burial, and explosion, which could lead to greater damage to life and property. Some studies have sought to find the main causes of fire accidents that can cause enormous damage to construction sites. One study shows that fire hazards in the category of unsafe site conditions are the leading cause of accidents on construction sites (Abdelhamid and Everett 2000). Unsafe conditions on construction sites include flammable materials stored on-site and hazardous construction activities that can directly cause fires. According to a related study, fires on construction sites are mainly caused by the transfer of sparks to the surrounding insulation during welding (Kim and Kim 2018; Lee 2012). In addition, various gases are stored on the site depending on the type of construction, which can increase the risk of fire and explosion on the construction site (Li et al. 2020). The location and method of storing combustible materials in construction sites was determined to be an important factor that can prevent fire accidents (Sielicka et al.). In this study, it was pointed out that site layout planning without adequate consideration of the storage location of flammable materials and heated devices are the main causes of construction site fires.

Some studies provide methods to evaluate fire hazards on construction sites, such as a fire hazard tracking system and an evaluation index (Hui et al. 2012; Kehui 2011). One study collected information to evaluate construction site fire safety, provided relevant indicators, and evaluated the fire safety on the site by calculating this mathematically. In addition, there is research to improve fire safety by providing a monitoring system that can detect fires on construction sites in real-time. In related research, a real-time construction site fire detection system was developed by collecting CCTV screen data provided from construction sites and developing an image-based fire recognition model (Su et al. 2021). However, these studies to evaluate or monitor the fire risk of a construction site are limited in regard to improving the fire safety of a construction site in the design stage of a construction project because the related data is collected after the actual construction site is started. If data related to fire on the construction site are collected in advance and the fire factors on the site are analyzed, it can be effective in providing regulation before the start of the construction project that will prevent fires. Also, when analyzing construction site fire factors, it is important to collect various data and analyze factors related to accidents on construction sites (Aires et al. 2010).

Web crawling application in construction research

Web crawling is a technique for systematically browsing the web for the purpose of web indexing (Paul et al. 2017). It is often used for tracking web documents on the Internet to effectively collect the information the user needs (Massimino 2016). Because the data on the web is huge, collecting web data manually can be time-consuming, and the accuracy can be reduced. However, web crawling technology automatically rotates the web server to repeatedly collect information that fits the purpose. These web crawling technologies are used in a variety of fields, especially in research involving decision models and prioritization (D'Haen et al. 2016; Guy et al. 2019). Recently, research on safety and security has been conducted through web crawling (Morgan et al. 2020).

The construction industry has also started using web crawling, and the most active field is construction material management and optimization. Related studies used web crawling technology to collect construction material information and provide automated processes. (Hong et al. 2019; Yang et al. 2018). Web crawling has also been used to manage massive documents in construction projects. As an example, web crawling was used to develop a system that collects the latest construction market text data and automatically assigns it to each applicable construction document (Moon et al. 2018). Recently, this technology has expanded into various fields related to construction. An example is a study that collects a variety of geographic information on the web and provides a model to predict air emissions from heating (Lopez-Aparicio et al. 2018). However, few studies have analyzed the factors related to the safety of a construction site using web crawling. In this study, web crawling technology was used to find factors related to site safety, which may suggest a new approach to improving construction safety.

Word embedding and network analysis

Word embedding is a technique that provides a way to express similar words with the same meaning through data analysis (Yin and Shen 2018). This is a new way to represent words and documents and is one of the key breakthroughs in deep learning (Lai et al. 2016; Yoshioka and

Dozono 2019). This method of analysis was mainly used as a new method for analyzing text or documents (Rekabsaz et al. 2017; Shao et al. 2017; Zhang et al. 2018). The main models using word embedding are Word2vec (Church 2017), GloVe (Hindocha et al. 2019), and fastText (Choi and Lee 2020), and Word2vec was used in this study. The Word2vec Model was created and published in 2013 by a Google research team. Word2vec has very efficient performance and accuracy (Ombabi et al. 2017). In addition, Word2vec has a lot of Google news pre-trained data (Khatua et al. 2019), so it is suitable for performing various analyzes based on pre-trained data.

Network analysis is one of the methodologies for finding the relationship between various types of raw data and consists of nodes and edges (Smith and Gorgoni 2018). Links between words can be expressed as structures, and relationships between words can be analyzed using them. Also, the correlation between nodes can be analyzed by calculating the Jaccard coefficient that can measure the similarity of sample data (Bag et al. 2019). This network analysis has been widely used in recent safety issues, such as pandemic research (Sandhu et al. 2016).

METHODOLOGY

In this study, media data related to fires on construction sites was efficiently collected using web crawling. To analyze the collected data, word embedding, and network analysis were used. Word embedding and network analysis play a role in providing relationships such as similarity by calculating vector values of important keywords in this study. In addition, Uniform Manifold Approximation and Projection (UMAP) was used to effectively express data analyzed by various methodologies in 2D space.

Web crawling

This study implemented a web crawling method for data collection. The web crawler usually traverses web pages by using a recursive algorithm and then goes over a certain range defined by researchers. The crawler stores data in a data structure that researchers can use efficiently for their studies (Mahto and Singh 2016). To begin with, the authors set the scope in which the crawler should travel. This research collected data from The New York Times, which is the top 3 media company in terms of newspapers by circulation and thus it has sufficient representative data for the study. In addition, some media companies restrict crawling or limit the amount of data for crawling while The New York Times has generous terms for crawling the data.

This research utilized the search term "construction accident" and retrieved the most relevant articles within 20 years. Among articles, the authors only handle data formatted in text. Since different formats such as blogs and interactive documents consist of irregular structures, these formatted articles cannot be analyzed. There are vital libraries for web crawling: Selenium, HTMLParser, and Beautiful Soup. Selenium is the set of tools that assists in the development of test automation for a web-based application (Mustika 2018). Selenium automatically traverses the web pages and stores data within the limit range. Next, HTMLParser was used for parsing HTML files in each article. With the parsed HTML, Beautiful Soup collected the data the researchers desired to utilize. Beautiful Soup is a Python package that analyzes HTML/XML, as it extracts and edits information on web pages. It provides a simple interface for building a text analysis prototype for text analysis and mining the data (Zheng et al. 2015). In this research, the authors retrieved the data from parsed HTML and sorted out the bodies, titles, and dates of the articles. The structure of the articles provided by The New York Times has a tag for body and title. Collected data were used for word embedding in the next part of this study.

Text data preprocessing

Before embedding the words, preprocessing is an essential phase to utilize text data. In this study, preprocessing is operated in three steps: Regularization, Tokenization, and Removing stop words. The study firstly regularizes all the articles by lowering the case. Natural Language Toolkit (NLTK) library was used to conduct the rest of the two phases Natural Language Process(NLP).

Using the NLTK library, the sentences were broken down into token-level, which breaks text into words, phrases, and symbols (Kannan et al. 2014). After sentences were tokenized, the authors removed “English stopwords” provided by NLTK. The definition of “English stopwords” is unnecessary words filtered before and after processing natural language data. Eliminating stop words reduces the dimensionality of vocabulary space, and examples of stop words are “the, in, a, an, etc.”(Vijayarani et al. 2015). Some studies involving NLP remove more words than the words in this study, but this study minimizes the target words. Given that word embedding requires sequences of words, removing numerous words interferes with the effectiveness of word embedding. Therefore, this research eliminates the minimum number of words.

Word embedding (Word2Vec)

Word2Vec is one of the most powerful methods to implement word embedding. Since word2vec assigns a vector to each word, word2vec shows advantages over other word analysis techniques. Before the presence of word embedding algorithms such as word2vec, researchers usually used one-hot encoding to analyze words. One hot encoding converts each word to specific (associated with) values(numbers) that differ from other words. Even though this helped them to conduct Natural Language Processing (NLP), the researchers could not answer the relation of each word because every word in one-hot encoding is independent. Word2vec resolves this problem, as it adds the embedding layer into the model. Word2vec follows a deep learning approach that probabilistically predicts word vectors by using a hidden layer. Since Word2Vec conducts unsupervised learning that is trained on raw text data, it creates word embedding by figuring out the maximum likelihood of word prediction from their context (Ling et al. 2015). The Word2Vec algorithm computes cosine similarity to find out similarity or dissimilarity between two vectors. In this study, cosine similarity plays a significant role in analyzing words, which helps to compare the words and construct relations among words.

There are two models for implementing Word2Vec: The Continuous Bag-of-Words model (CBOW) and the skip-gram model. Even though two models are both widely used in conducting word2vec, these models can be optimized depending goals and direction of the research. The skip-gram model computes the possibility of the target words in several contexts from a word. The model has multinomial distributions for the outputs. It can be effective when the research tries to figure out fewer frequency words. On the other hand, CBOW is more suitable for studies that concentrate on high-frequency terms. In this study, since the authors focused on important keywords that can explain accidents on construction sites and frequently appearing words, the authors implemented Word2Vec with the CBOW model. The CBOW model predicts words by considering context, which means the target word is predicted from surrounding words (W(t-n), W(t-n+1)...W(t+n)) of the target word, and thus it has one output vector. Input consists of on-hot encoded input context words. Figure 1 shows an overview of the CBOW model

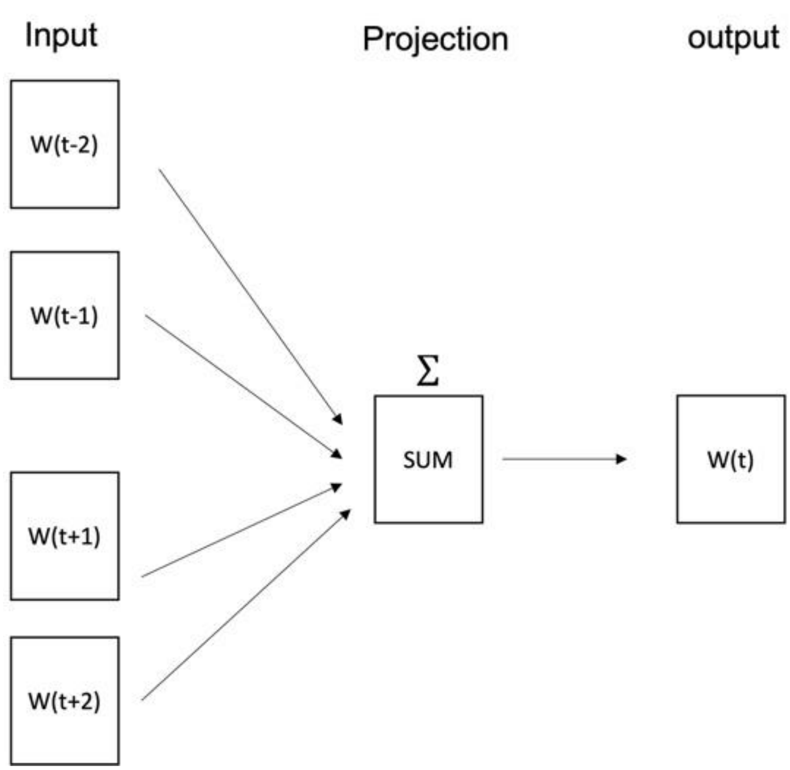

Figure 2-1 Overview of the CBOW model

In the projection phase, input generates output through a hidden layer $\boldsymbol{h}$,and it is computed as follows,

$$\boldsymbol{h} = \frac{1}{C}\mathbb{W} \cdot (\sum_{i=1}^{C} x_i)$$

, where C is window size, x is input and $\mathbb{W}$ is weighted matrix. Output layer(u_j) can compute the possibilities ,and output(y_j) is created by through the soft-max function.

$$y_j = p(w_j \mid w_1, \dots, w_C) = \frac{exp(u_j)}{\sum_{j'}^{V} exp(u_j')}$$

, $where\ V\ is\ vocabulary\ size$. Through this process, the CBOW model returns the one output from surrounding words. There are several parameters for conducting word2vec with CBOW. The authors restricted the minimum count of words to 200. In other words, the authors only considered words that appeared at least 200 times in overall articles. This model set the words vector into 400 dimensions, which represents the better accuracy of the model. The window size was set to 5.

Network analysis

Networks are a versatile method to show and analyze simple and complex interactions among factors in articles and thus they are used for studies in diverse areas. The network representation is simple but rigid since many parts of a specific system are sorted out and concentrate on the interaction among its elements (Menczer et al. 2020). The network is represented with nodes and edges that connect nodes to each other. In this study, the top five frequent keywords and similar words that have over 0.5 cosine similarity with the keywords are represented as nodes. The authors measured the Jaccard coefficient between each keyword to compare similarity and dissimilarity. Jaccard coefficient calculates the result of the division between the number of features that can be seen and the total number of features (Niwattanakul et al. 2013). The authors examined the interaction over five keywords and computed the Jaccard coefficient between two keywords over ten combinations of keyword sets.

Uniform Manifold Approximation and Projection (UMAP)

In this study, firefighters were asked about the effect of building cycle on fire and fire suppression. Eighty-two percent of the respondents said that the building cycle had a significant effect on firefighting activities. In addition, more than 76% of the respondents said that the progress of construction has a great impact on firefighting activities. The answers to the questions about the most difficult times during the building cycle were answered during the occupancy (57.1%), during the construction (28.2%), and the no occupancy for demolition (14.7%). During occupancy (73.1%), the highest response rate was found in the question of the stage of the most likely fire during the building cycle. Next, During Construction showed a response rate of about 20%.

With the result of the Word2Vec, the authors generated 400-dimensional word vectors. UMAP was used to visualize the vector into a low dimension space as UMAP is the state-of-the-art technique for dimensional reduction. Dimension reduction creates low dimensional space without loss of structure in high dimensional space. UMAP has been widely used in various fields with larger sizes of data (McInnes et al. 2018). Riemannian geometry and algebraic topology are the theoretical ground that constructs UMAP. UMAP operates on weighted graphs, and it uses k-neighbors to cluster groups. UMAP is usually compared with the alternative dimensional reduction technique, t-SNE. Compared with t-SNE, UMAP significantly performs faster and more efficiently as well as better preserves global structure. The t-SNE more easily suffers from the curse of dimensionality on the large-scale data set(McInnes et al. 2018). However, UMAP shows almost no restriction on the embedding dimension, which leads to being feasible for deep learning. Therefore, this study conducted dimensional reduction with UMAP ,and visualized word vectors in a two-dimensional space. Every word expressed in a multi-dimensional vector was reduced to a two-dimensional vector, and then the research team annotated every dot in a two-dimensional graph. Since the authors focused on five of the most frequent keywords, the annotation of keywords had larger fonts and similar words with each keyword represented in the same color.

RESULTS

The results section of this study consists of 6 subsections. The detailed and statistical approach to the data collected by web crawling is described in the preliminary analysis and basic statistical analysis sections. Based on the collected data, the main keywords of this study were analyzed using word embedding and network analysis methodologies. In order to intuitively provide data analyzed in multiple dimensions, this study presented results in 2D space using the UMAP concept. Finally, the collected data related to fire accidents on construction sites were extracted and analyzed, and specific factors related to the fire accident on the construction site were provided.

Preliminary analysis

Through web crawling, a total of 1,010 relevant articles from the New York Times was found. "Construction accident" was used as a search term on the New York Times website, and the top 1,010 relevant articles were retrieved. Of all the data collected by web crawling, only document type articles were used for analysis, and 149 articles of the interactive document and blog type were excluded from the analysis. The article types excluded from the analysis are not valid for scraping because of the irregular structure. Therefore, a total of 861 articles were analyzed. After completing data cleaning, the authors classified articles according to composition. It was confirmed that the articles are generally composed of title, date, and body, and the text data included in the article was classified in consideration of the compositions. The Beautiful Soup library was used to classify and scrape the parts needed for analysis. Table 2-1 shows the collection of source data.

Table 2-1 Source data information

Category	Information
Issued duration	2000.01.01 – 2019.12.31
News source	The New York Times
Total number of articles	1010
The number of relevant articles	861
Total number of words	453,283

Basic statistical analysis

The five keywords were “fire”, “fell”, “collapsed”, “building”, and “people”. The selection criteria considered the types of accidents on construction sites and keywords representing building fire. In many studies, three major elements of building fire are defined as a “fire”, “building”, and “people”(Kim et al. 2021). In addition, “fell” and “collapsed” are the most frequent types of accidents on construction sites (Abdullah and Wern 2011).

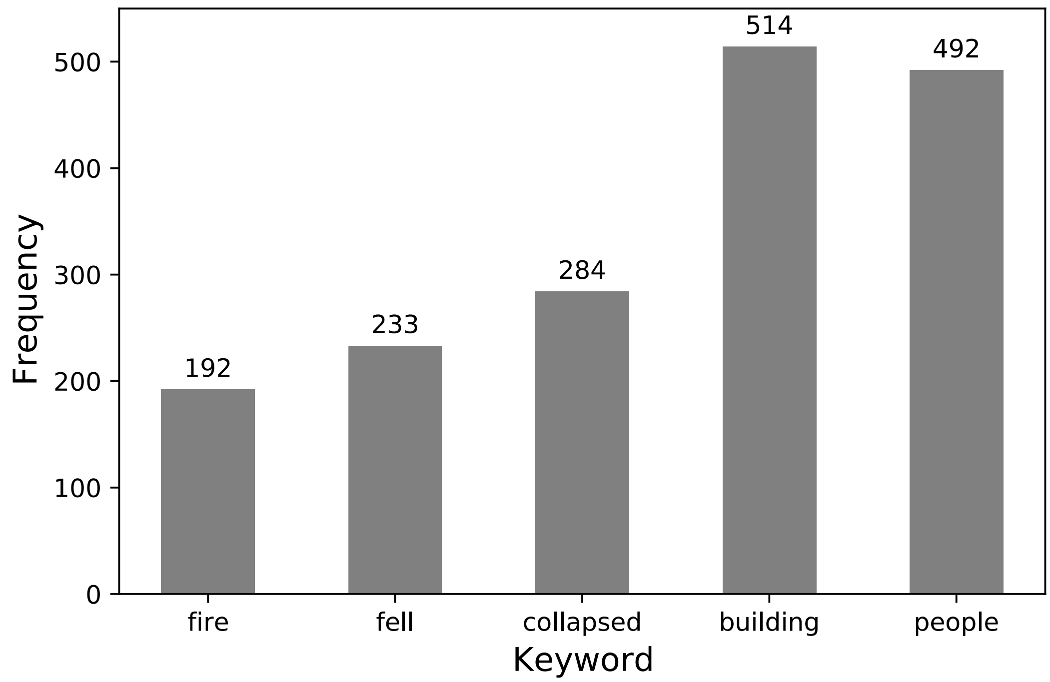

Figure 2-2 The media exposure level of keywords

According to the analyzed results, buildings and people were the most frequent. The "fire" "fell", and "collapsed" keywords related to the type of accident on the construction site showed a relatively similar level of media exposure. Since "fell" and "collapsed" are similar words, it may be reasonable to combine them when comparing with the keyword of "fire". After combining, the "fire" has a media exposure level of 27%, and "fell" and "collapsed" combined have a level of 73%. This result differs significantly from the Bureau of Labor Statistics (BLS) analysis of construction site accident frequency. According to the BLS report, among the accidents on the construction site, fell and collapsed accidents account for about 40% of all accidents, and fire accidents are 2%. When converted to 100%, fire accidents have a frequency of 5% in fall-related accidents. There is a large gap in the media exposure level for "fire" presented in this study and the frequency of fire accidents in the BLS report. The possible reasoning for this difference can be explained by the characteristics of the media that the authors described in the introduction section. Fire-related accidents may be exposed to the media more than fall-related accidents, which may explain that fire accidents on construction sites have a greater severity than fall-related accidents.

Word embedding in vector space through Word2Vec

In this study, cosine similarity was used to provide semantic similarities between keywords. A combination of ten pairs of key words were generated and the similarity of each pair of keywords was calculated, as shown in Table 2-2. To confirm the reliability of this study, the cosine similarity of “fell” and “collapsed”, which are the most similar words among the analyzed words, was first checked. The similarity between the two words was 0.951, which was calculated much higher than the similarity between the other keywords. Through this, the authors can confirm the reliability of this study. In the results related to the “fire” keyword, the “fire” keyword was more similar to the “building” keyword than the “people” keyword. The similarity between “fire” and “building” was 0.525, and the similarity between “fire” and “people” was 0.331. It can be interpreted that “fire” showed a much higher degree of similarity to words related to the “building” than words related to “people”. For the “fell” and “collapsed” keywords, their similarities with “people” and “building” were very close. Specifically, “fell” has a similarity with “building” and “people” of 0.443 and 0.486, respectively. Also, the similarity between “people” and “building” was calculated as a negative value, indicating that the correlation between the two keywords is not great.

Table 2-2 Similarity between keywords

Relationship between keywords	Similarity
Fire and Building	0.525
Fire and People	0.331
Fire and Collapsed	0.691
Fire and Fell	0.728
People and Fell	0.486
People and Collapsed	0.510
People and Building	- 0.144
Building and Collapsed	0.554
Building and Fell	0.443
Fell and Collapsed	0.951

Table 2-3 shows the top 20 similar words to each of the five selected keywords. Among the similar keywords related to “fire”, “fell”, and “collapsed”, “Monday” and “Friday” showed across the board. Since “fire”, “fell”, and “collapsed” are all closely related to the types of accidents, this result is very consistent with the previous research findings on the construction industry's 'The distribution of injuries' across the weekdays (Wigglesworth 2006). However, the “fire” keyword showed a high degree of similarity with the word “night”, which was not observed in “fell” and “collapsed” keywords. Generally, the possibility of spread increases when a fire occurs at night. There are many words that have the meaning of “administration” or “inspection” in words with high similarity to the “building” keyword. The top five words with the most similarity to the “building” keyword are “department”, “inspectors”, “issued”, “commissioner”, and “city”, and these words tend to have a common meaning. Also, most of the words with a high similarity to the “people” keyword are related to workers or activities related to construction workers.

Table 2-3 Top 20 list of similar words of keywords

Fire	Fell	Collapsed	Building	People
police	collapsed	fell	department	died
investigators	side	debris	inspectors	injured
authorities	floor	floor	issued	killed
hospital	debris	ground	commissioner	men
injuries	ground	west	city	accident
chief	street	steel	cranes	authorities
night	Friday	street	contractor	debris
floor	avenue	death	violations	train
man	death	side	tower	injuries
dead	morning	site	crane	dead
Monday	worker	Friday	investigation	workers
Friday	injuries	march	office	ground
driver	Monday	wall	equipment	Monday
investigation	injured	injuries	site	crash
truck	dead	authorities	district	march
officials	wall	Monday	charges	residents
death	authorities	avenue	investigators	members
worker	steel	tower	company	Tuesday
fell	crash	Tuesday	officials	Thursday
working	died	injured	mayor	cars

Network analysis of keywords

Network analysis is an analysis method that describes the relationship of data with nodes and edges. By using this network analysis, not only is it efficient to interpret the relationship between nodes of data, but it is also able to retrieve the node that has an important impact on the network other than the main nodes. In this study, keywords are nodes, and words with a cosine similarity of 0.5 or higher are connected by edges. Nodes depict five keywords and similar words, and the number of nodes is 136. Edges represent connection among words, there are 353 edges in the network. In addition, the Jaccard coefficient between each keyword was calculated and provided. The Jaccard coefficient values between keywords are shown in Table 2-4. The Jaccard coefficient is a statistical value used to measure the relationship and diversity of sample data. Through this, the network of each keyword can be expressed as one unified network. This network graph is visualized in Figure 2-3. This entire network has different sizes for nodes and annotations based on degree. This means that the more nodes are connected (the higher degree), the larger the size of the node. Among the keywords in this study, “collapsed” has the largest degree. In addition to the five keywords, the word in which the node size is noticeably larger is “death”. This shows that the word “death” has a high degree besides the five keywords in the network in Figure 2-3.

Table 2-4 Jaccard Coefficient between keywords

Relationship between keywords	Jaccard Coefficient
Fire and Building	0.262
Fire and People	0.262
Fire and Collapsed	0.513
Fire and Fell	0.541
People and Fell	0.484

People and Collapsed	0.424
People and Building	0.020
Building and Collapsed	0.205
Building and Fell	0.160
Fell and Collapsed	0.772

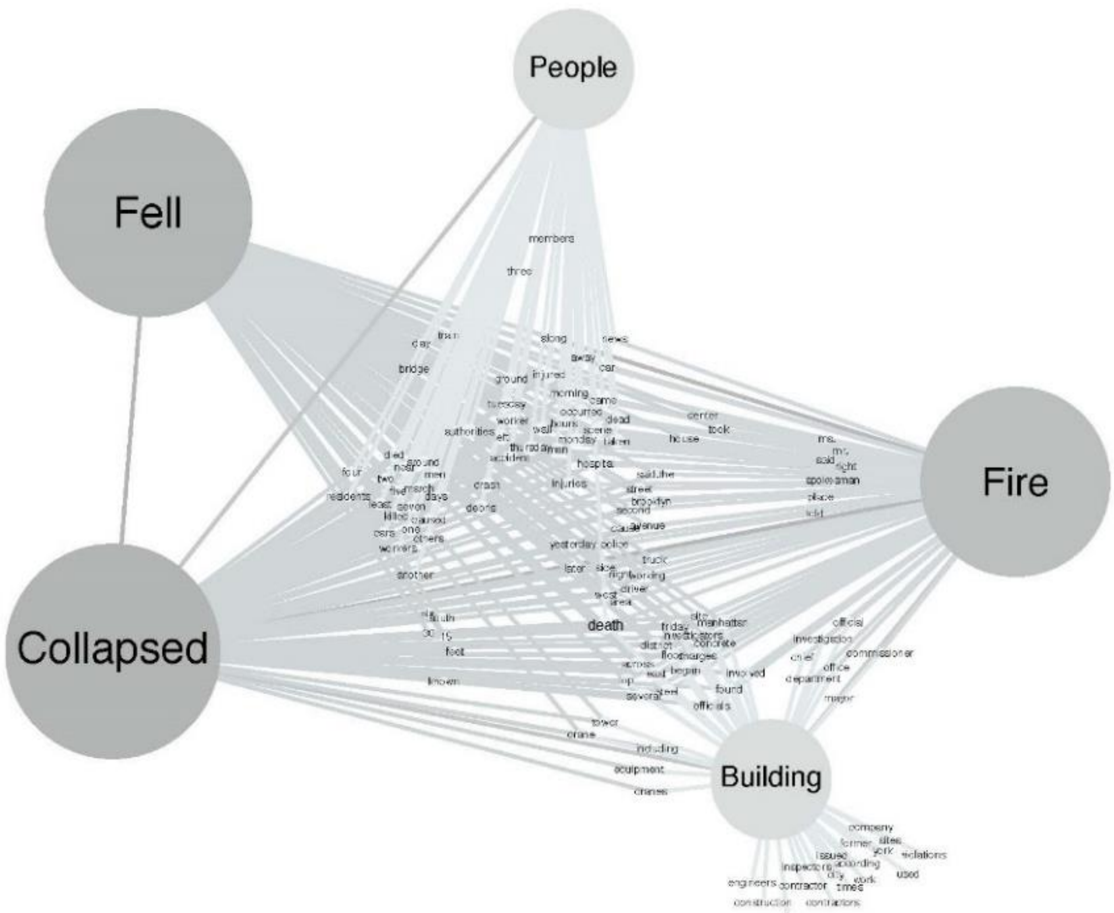

Figure 2-3 The network of keyword's similar words

Visualizing with UMAP

In this study, the results analyzed through Word2vec were visualized in two-dimensional space by using UMAP. It enables seeing the sparsity among keywords and their similar words. Figure 4 shows the overall UMAP graph for this study. In order to increase the discernment of the graph, each keyword is expressed in a different color, and the range of each keyword is indicated by gradation. The gray dots on the UMAP are words with low similarity to keywords. As shown in the figure below, “fell” and “collapsed” were almost overlapped on the UMAP. This means that the similarity between the two keywords in UMAP is very high. In the case of the fire keyword, the range appears wider than other keywords, and there is an intersection with the “building” keyword. For the “fell” and “collapsed” keywords, the “people” keyword appeared to be closer than the “building” keyword. By using this, this study can provide information about the relation among words in the two-dimension plane according to the similarity of words.

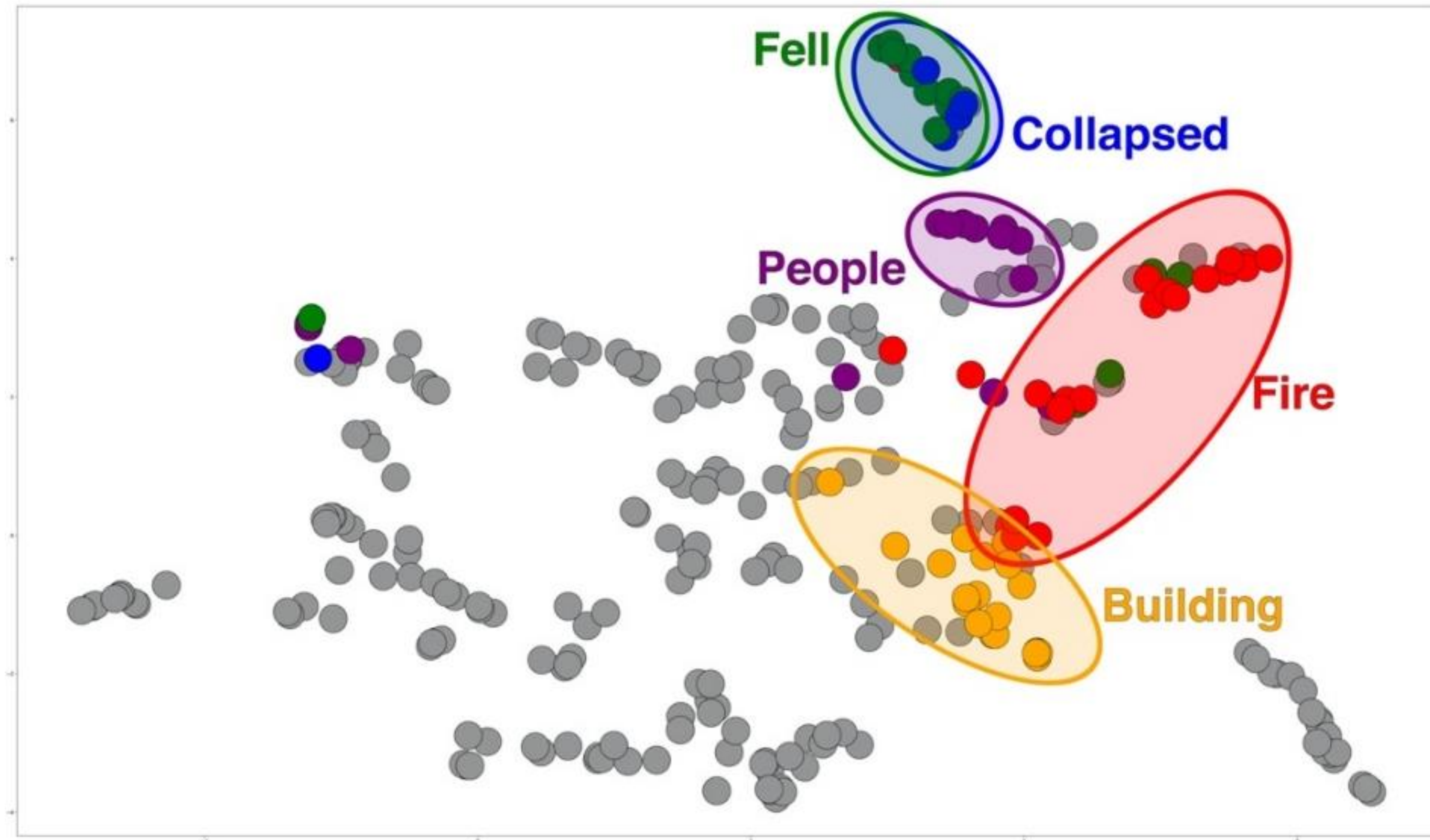

Figure 2-4 Geometric representation of keywords in 2D space through UMAP

Factors related to fire accidents

Articles related to fire accidents were classified to provide an in-depth analysis of the factors related to fire accidents on construction sites. Through the classification of the season in which the article was written, most news articles related to fire accidents on construction sites were written in summer. According to the BLS report, which investigates the distribution of injuries throughout the year in workplaces, more injuries occur during the summer season than at other times of the year (Pierce 2013). In addition, the two sources had a common feature that injuries were much less frequently reported near the end of the calendar year. Table 2-5 shows the distribution of articles and injuries related to fire accidents on construction sites by season.

Table 2-5 Distribution of articles related to fire accidents and injuries by season

Division (season)	Collected articles (N=125)	The Bureau of Labor Statistics
Spring (Mar-May)	29.6%	24.8%
Summer (Jun-Aug)	31.2%	27.4%
Fall (Sep-Nov)	16.0%	24.2%
Winter (Dec-Feb)	23.2%	23.6%

In this study, the topics and major factors of articles dealing with fire accidents on construction sites were classified. In the collected articles, the biggest factor in the fire accidents was explosion caused by a chemical gas leak. In addition, building-related factors such as building code violations, lack of regular inspection, and inadequate fire safety systems account for about 40% of major factors in fire accidents. Factors related to the construction activities on sites, such as welding and activities related to demolition and renovation, accounted for a relatively small proportion. It can be confirmed that these results are consistent with the results of the analysis through Word2vec in this study. Table 2-6 shows the main factors in the articles dealing with fire accidents on construction sites.

Table 2-6 Distribution of major factors in fire accidents

Factor	Percentage
Explosion related to chemical gas	20.0%
Violation of building and fire code	16.8%
Lack of building and site inspection	11.2%
Inappropriate fire safety system	10.4%
Carelessness	8.8%
High wind	5.6%
Absence of an evacuation plan	5.6%
Activities related to demolition	5.6%
Welding	4.8%
Activities related to renovation	3.2%
Etc.	8.0%

DISCUSSION

This study explored the relationship between factors related to accidents on construction sites through web crawling and deep learning approaches. It is interesting to find that the media exposure level for fire is disproportionally higher compared with the typical construction Fatal Four. The difference between the frequency of actual accidents and the level of media exposure can be used as evidence to confirm the great impact that fire accidents have on construction sites. As OSHA publishes reports and statistics related to the 'Fatal Four' every year, efforts in addressing them have been the focus of many stakeholders in the construction industry. Fire accidents have a low frequency, and thus fire accidents are excluded from the list of major accidents on the construction site. However, the results of this study confirmed that fire accidents have a higher media exposure level than other types of accidents that happen more often on job sites, which signifies the greater social impact and severity than other accident types. This result is in line with the fire accident characteristics. Fire accidents are more likely to lead to secondary accidents than other types of accidents on the construction site. Due to the spread of fire, it can affect the surrounding buildings and roads. Therefore, fire safety should at least be equally emphasized when developing on-site safety regulations and policies.

Construction site safety training period

This study conducted an analysis using Word2vec, one of the word-embedding models. The results through word-embedding can be used as useful data to explore ways to improve the safety of construction sites. One example is that the results of this study can be used to determine the safety training cycle of a construction site. The list of words with high similarity to the three keywords related to accidents ("fire", "fell", "collapsed") commonly includes "Monday" and "Friday". This result is consistent with statistical data related to accidents on construction sites. According to a related study, workers' injuries on construction sites were the highest on Monday (Wigglesworth 2006). The common results of these studies can be helpful in determining the timing of worker safety training to improve fire safety on construction sites. Fire accidents on construction sites can be effectively prevented by conducting fire safety training for the workers and inspections on Monday, which is the start of the week of the construction project. In addition, since the results of this study and the statistical data provided by The Bureau of Labor Statistics show that the frequency of accidents on construction sites increases in summer, this result should be reflected in the annual safety training schedule.

Fire detection system on the construction site

The word "night" has a high degree of similarity in the "fire" keyword. Construction sites tend to have few occupants after work hours compared to other building types. In particular, a construction site may have a limited number of employees staying overnight for monitoring or may be completely empty, which does not allow for early detection of a fire on a construction site. Also, during the construction phase, it is more difficult to recognize fires because safety equipment such as fire and smoke alarms have not been completed (Hamid et al. 2003). A report published by the National Fire Protection Association that breaks down direct property damage from construction site fires by time confirms this risk. According to this report, direct damage caused by fires on construction sites occurred from midnight to 4 am, accounting for 31% of the total damage (Campbell 2017). This is about a 15% increase in direct damage compared to other time periods.

Early fire and smoke detection is very essential to the success of fire safety on building construction sites. Safety regulations and policies should be augmented so that fire and smoke detections system or monitoring system should be required on construction sites to detect night-time fires on construction sites. In addition, there is a need for a plan to expand the periodic fire safety system inspection program that is conducted on completed buildings to the construction stage. Fire protection systems on construction sites are more likely to be exposed to the risk of damage compared to completed buildings, and these risks create a possibility that fire protection systems could not function properly in the event of a fire at night. According to related studies, smoke and fire sensors are difficult to use properly in the construction stage due to the open environmental conditions and environmental complexities of the construction site (Su et al. 2021). These risks are also consistent with the finding of this study. It was found that there were many words that had a meaning of inspection with a high similarity to the "building" keyword.

These results show that periodic and reliable on-site inspections are as important as installing additional fire protection systems to improve fire safety on construction sites.

Fire safety regulations on the construction site

According to the results of this study, the “fire” keyword showed higher similarity to the “building” keyword than the “people” keyword. This can be explained by the fact that fires on construction sites have a higher relationship with building factors. The violation of the building code and the lack of regular inspection and training appeared as the main contributors to fire on the building construction site. According to the collected articles, the most noted violations on the construction site were dangerous working conditions such as the absence of a fire safety system and inadequate safety supervision. Improper use of fire extinguishers, such as those used to prop open doors, and exposure of electrical cords that could cause a fire have also been found in violations of the regulations. In addition, the absence of illuminated signs that can be found in the dark that affects workers' evacuation in case of fire on a construction site was pointed out. This result provides important evidence that many fire accidents can be sufficiently prevented through site regulations and behavior changes that reflect the building factors of the construction site. Factors related to the fire safety system on construction sites can improve safety at the design stage of construction projects, and dangerous working conditions on the site can be prevented through appropriate safety supervision and inspection by the site manager.

The current construction site fire safety system does not sufficiently consider the characteristics of the construction stage. Therefore, a blanket approach to address all stages of construction using the same regulation and fire safety system might not be appropriate. It is necessary to consider applying differentiated fire safety regulations for each construction stage by defining major factors and related activities at each stage of construction. An alternative is to determine the fire risk in related activities based on the building code and install a separate fire safety system at the location and time vulnerable to fire on the construction site. For example, some of the news articles reported that site management was aware that welding activities were the main cause of fires on construction sites, but no separate fire safety system was considered for welding activities on site. Indiscriminate exposure of electrical cords and storage of combustible materials on the construction site must be thoroughly managed during welding. In addition, under the current construction safety program, it is difficult to efficiently prevent fire accidents due to the lack of OSHA's fire training program at the construction stage and related research to improve fire safety on construction sites. Therefore, more research and investment in fire safety on construction sites are needed.

Risk of explosion

In the case of explosions that account for the largest proportion of fire accident-related articles collected in this study, additional modifications to the fire-related regulations and site safety policies will be required. There are many activities where chemical gas is used on construction sites. Depending on the progress of the construction project, there is a possibility that gas pipes can be involved, and safe control on gas pipes is an important factor in improving fire safety. In particular, construction activities that use gas in confined spaces, such as underground, increase the likelihood of fires and explosions on construction sites. Since gas leaks and explosions in confined spaces can cause greater injuries to workers, additional fire safety systems and protective equipment are required through additional regulations based on construction activities. It is possible to consider a method that can inform workers of the danger of explosion in an enclosed space by applying a portable sensor that has been recently used for indoor air quality analysis.

However, many articles collected in this study point out that there are still no major regulatory changes to address explosion risks. Each regulation should include a manual on handling hazardous gases for each construction activity, and external conditions such as weather that may affect gas expansion should also be considered. In addition, all stakeholders involved in the construction site must be continuously provided with detailed information on where and how much flammable materials such as chemical gases are stored. It is essential to modify and strengthen the building and fire regulations related to the management and inspection of chemical gases according to the construction stage.

CONCLUSION AND RECOMMENDATIONS

Researchers have rarely focused on fire safety on construction sites due to its low occurrence compared with other accidents, such as falls. Researchers also have not adequately explored the factors of fire accidents on construction sites and the effects of fire accidents. This study contributes to the body of knowledge by providing evidence to support the importance of fire safety, identifying specific causes and related factors of fire accidents, and providing specific recommendations to address fire safety on construction sites.

To analyze the influence of factors related to fire accidents on construction sites, new approaches such as web crawling and word embedding using deep learning were introduced and used in this study. This study provides some evidence of factors that influence construction site fires based on previously unstudied media data. It was found that fire accidents could be an important factor threatening the safety of construction sites. The results of the current study suggest that when developing regulations and policies to improve construction safety, the risk of fire accidents should not be overlooked and should be equally considered. In addition, this study suggested that building and fire code violations, lack of regular inspections, and an incomplete fire safety system were major factors in fire accidents on construction sites. This is not a result of the careless and risky behavior of workers on construction sites, but the factors that can be prevented with appropriate regulations and on-site protection systems. By finding the reasons for fire accidents on the construction site and exploring the relationship between the related factors, this study makes meaningful contributions to developing safety-related regulations that consider the characteristics of construction sites.

The methodology and results chosen for the current study presented several limitations to the analysis. The target of the data collected in the current study is limited, analysis of other targets that can investigate the impact of fire accidents, such as data related to fire insurance on construction sites, is required in addition to the news article. Future studies should analyze factors related to fire safety on construction sites from various perspectives based on data collected from different targets. In this study, an analysis was conducted to find the relationship and similarity between representative keywords, but analysis of sub-words may be required based on sufficient data collection. The current study analyzed representative keywords such as "fire" and "building", but in future studies, the relationship between the sub-keywords belonging to the representative keywords can be explored. Therefore, future studies can analyze the factors related to the fire safety of construction sites by subdividing them, and through this, fire safety on construction sites can be considered on a larger scale. Applying more advanced data collection and analysis techniques to improve the reliability reinforces the results of this paper. New methodologies in other fields, such as deep learning, must be actively introduced to suggest ways to improve the fire safety of construction sites.

CHAPTER III

EXAMINE FACTORS INFLUENCING FIRE SAFETY ON BUILDING CONSTRUCTION SITES: A FIRE OFFICER'S PERSPECTIVE

ABSTRACT

Fire safety on building construction sites as an important part of on-site safety management has been rarely studied for construction sites. To fill the knowledge gap, this research was conducted to identify important influencing factors on fire safety on construction sites as well as their relationships through a questionnaire administered to 1,492 fire officers in South Korea. The authors divided the questions into three factors (building, fire, and human) to distinguish factors affecting construction site fire safety. For the majority of the factors, there was high agreement on the importance level of factors affecting fire safety during construction between administrative and field firefighters. However, perception disparity did exist. The authors identified the interrelationships between factors through path analysis and structural equation modeling. The paths of factors showed the highest direct relation between building and fire factors and that the human factor played a mediator role between human and fire factors. The study contributes to the overall body of knowledge by providing officers' insights into the factors contributing to fire safety on construction sites as well as the quantification of the interrelationship among the factors, which will serve as the basis to develop effective safety measures, such as policies and regulations at both the project level and industry level to address fire safety on construction sites. The SEM model developed in this study can also be used as a base model for future related studies.

INTRODUCTION

In recent years, various skyscrapers and huge construction projects are being planned or are under construction in populous urban areas (Hardison et al. 2014). As construction sites get taller and larger, the possibility of fire and traffic volume for evacuation increases when incidents occur (De-Ching et al. 2011; Hinze and Raboud 1988; Mohan and Zech 2005). In addition, many construction projects are required to compress their schedules due to clients' pursuit of monetary benefits associated with an earlier start for their business operations. When a construction company is working under a crashed schedule, it is prone to neglect certain safety issues. As a result, various accidents, including fire accidents, occur on construction sites.

Most fire safety studies have limitations that focus on the building in use, not the building under construction. The building under construction has different characteristics from the completed building according to the construction schedule, such as lack of a fire safety system, lack of entrance guidance, and storage of flammable materials. This is especially dangerous on construction sites because fire safety equipment such as smoke detection systems and sprinklers may not be completed depending on the progress of the construction project (Hamid et al. 2003). Fire accidents on construction sites are highly likely to cause secondary accidents such as collapse, burial, and explosion (Yang et al. 2018). In addition, fire accidents on construction sites are greatly affected by the external environment, such as hazardous materials, weather, and surrounding buildings, making it difficult to control and prevent them. Also, in order to assess and manage the level of fire safety on construction sites, most researchers use computer simulation software (Dagan and Isaac 2015; Ingason et al. 2010; Li et al. 2016).

However, studies on fire safety on construction sites rarely consider fire officers' perceptions. Therefore, to enhance fire safety on construction sites, this research conducted a questionnaire related to fires on construction sites among fire officers. Based on the survey of 1,492 fire officers in Korea regarding their perceptions on construction site fire accidents, this paper analyzed factors that are important to reduce fire accidents on construction sites. This research used path analysis and structural equation modeling to analyze the interaction between factors and identify critical factors for enhancing fire safety on construction sites. The results of this study can also be used as the basis to evaluate the efficacy of the current fire safety regulations and measures on construction sites.

BACKGROUND AND LITERATURE REVIEW

Twenty-five to 40% of workplace accidents all around the world are construction site accidents (Aires et al. 2010). Accidents on construction sites not only cause a sizeable number of casualties but also negatively impact the cost of the projects and project duration (Hosseinian and Torghabeh 2012; Zou and Zhang 2009). In particular, fire accidents on construction sites can cause secondary accidents such as collapse and explosion, which can significantly increase the cost and duration of the project. To prevent fire accidents on construction sites, improving the safety perception of construction workers can play an important role (Abdullah and Wern 2011; Cooper and Phillips 2004; Hallowell and Gambatese 2009; Zou and Zhang 2009). In order to develop effective safety programs to improve workers' fire safety awareness on the construction site, an adequate understanding of safety factors affecting construction fire safety is necessary (Hallowell and Gambatese 2009; Kines et al. 2010; McDonald et al. 2009; O'Toole 2002). To apply various methods to improve fire safety on the construction site, an exploration of the basic elements of building fire should be preceded.

Elements of building fire

In general, elements related to building fire are defined as a building, fire, and human factors (Baozhi 2001; Wolski et al. 2000). The first factor to have a direct influence on the building fire is the nature of the fire (Kobes et al. 2010). Fire refers to the ignition and combustion process of substances that produce heat and smoke. This also includes flammable substances that enable ignition in buildings. Building factors have a direct impact on the fire since a building is a physical space where people are present and various activities take place. Fire safety systems such as sprinklers and fire detectors, which are important factors in the initial suppression of building fires, fall into this category. Finally, the human factor includes all personnel residents in the building and all personnel responding to it in the event of a building fire. Human nature can have a direct and significant impact on the degree of fire response performance (Kuligowski 2016). Also, these three elements determine the level of fire response performance in the event of a fire in a building (Kobes et al. 2010).

While many studies define the three elements of a building fire, there is an insufficient discussion about whether these three factors affect each other. In a study published in 2001, discussions began on the relationship between fire and humans, but no studies have been found on the effects of human-related factors on building fires (Sime 2001). There is a limitation in analyzing building fires only from the building side or only from the human side. It is important to study these three factors independently, but it is also necessary to explore the relationship between the three factors by analyzing the effects of these factors on each other. In addition, related studies define the building factor based on the completed building, but in order to consider fire safety on the construction site, it is necessary to extend the building factor to the construction stage. These three factors can directly or indirectly affect the fire, and when proper interventions revolving around these factors are properly introduced, fire accidents on the construction site can be managed efficiently.

Fire risks on construction sites

Various safety threats, such as dangerous activities, lack of safety awareness, violation of safety regulations, and lack of safety-related systems on construction sites can lead to fires on the construction site (Maluk et al. 2017; Wolski et al. 2000). There are related studies investigating the main causes of fire accidents on construction sites. According to Lee (2012), fire on construction sites is mainly caused by sparks transferred to the surrounding insulation during welding. Likewise, Kim and Kim (2018) found that most fire accidents occurred due to sparks by analyzing big data related to fire accidents during construction over the past ten years. To determine the main factor of threats to construction site safety including fire, some studies used not only the statistical approach but also the perception or behaviors of construction workers and managers through surveys (Abbas et al. 2018; Chan 2011; Shin et al. 2014; Zou and Zhang 2009). Abbas, Mneymneh et al. (2018) interviewed construction workers to investigate their fire risk perception about dangerous activities and the study concluded that the workers perceived unprotected openings, fire, and welding as the most dangerous factors. According to Shin, Lee et al. (2014), fire safety during construction could be affected by the unsafe acts of workers as well as unsafe conditions on construction sites such as flammable material.

However, since survey-related studies targeting construction workers are based on case studies and projects, they are limited in application to specific construction sites (Abbas et al. 2018; Shin et al. 2014). To overcome these limitations, it is an important approach to collect various data and analyze factors related to accidents on construction sites. In one study, the history of EU construction safety policies was collected and the effects of these policies on the accident rate on construction sites were analyzed (Aires et al. 2010). According to the study, it is important to continue to reflect the opinions of official bodies and other affected parties in order for new policies to improve safety on construction sites to be effective. Another study has collected literature data on human behavior related to fire safety in buildings, suggesting a new approach to the fire safety design of buildings (Kobes et al. 2010). This study emphasized the need to consider human psychology and characteristics beyond traditional building fire-related studies and mentioned that the starting point of fire prevention measures on construction sites should be an understanding of the interaction between human factors and the characteristics of a building. To effectively control a fire accident on a construction site where many factors have a complex effect, the traditional approach to building fire safety will have to be supplemented by various analysis approaches from the field

Point of departure

Fire accidents on the construction site can cause serious damage not only to the workers but also to those who reside nearby or pass by the site. However, there is little research on fire accidents during construction (Ingason et al. 2010; Jeon and Hong 2007). There is no proper standard on firefighting for buildings under construction so that many workers are directly exposed to fire risks (Jeong et al. 2014). In addition, most of the related studies only focused on the causes of fire accidents and the behavior of managers or workers. The construction project managers and workers are construction experts but may still lack sufficient knowledge about fire safety. Therefore, in order to have an in-depth analysis of fire safety on construction sites, the opinions of the group with expertise in fires such as fire officers are important as well.

Therefore, this study analyzed the perception of fire officers who have extensive experience and specialized knowledge in firefighting and suppression of fire accidents on construction sites. In this study, the survey was conducted in South Korea to investigate the factors affecting building construction site fire accidents from the fire officers' perception. Based on the collected data from the survey, path analysis and structural equation modeling were used to examine the relationship between factors affecting the construction fire safety. This study contributes to the overall body of knowledge by providing a direct assessment of the importance of factors determining fire incident results on construction sites from the perspective of fire officers as well as the perception disparity among the firefighter groups that need to be balanced in regulation and policymaking decisions. The identified relationships and paths identified between the different factors can serve as the basis to develop intervention measures to better improve fire safety in construction.

METHODOLOGY

Data collection procedures

Based on several previous studies related to the importance of building life cycle in firefighting, the questionnaire was constructed to reflect the situation of Korean fire officers. After reviews by various fire officers in Korea, questionnaires were revised reflecting the culture and situation of

Korean f fire officers. South Korean fire officers were invited via e-mail to participate, without incentives, in an online survey administered in spring 2018; snowball sampling was used to get as many fire officers as possible to participate. To request the contacts of potential participants in South Korea, the research team contacted fire department representatives and fire officers' associations in each region via e-mail. This survey consisted of a total of 19 items; 4 items were used for measuring the importance of building life cycle in firefighting and 15 items were used for measuring how important they actually are in determining fire incident results. The Institutional review board (IRB) approval (No. EN-18-3) for this research was obtained before data collection. Any incomplete responses with missing data were deleted, leaving 1,607 responses. Additionally, 115 outliers were eliminated by using the z-score for all survey items, thus the full sample size was 1,492. In this research, the authors surveyed Korean fire officers in order to investigate the factors influencing fire on the construction site. The demographic data were shown in Table 3-1. One thousand four hundred and thirty-one (94.7%) of the fire officers were male and 78 (5.3%) of the fire officers were female. According to the age variable, the majority age of the sample was between 30 and 39 (35.8%). In terms of final degree, the survey respondents most frequently answered 'BS equivalent' and 'Associate degree or 2-year college equivalent'. Among the survey participants, 53.2% of the respondents worked in the fire department for more than 10 years, and 30.8% of the respondents worked for more than 21 years. Respondents with less than a year of experience were 8.6%, and overall, it was found that employees with sufficient experience in the fire department participated in the survey. Based on the job specialty, firefighting work is the largest category (46.7%), followed by ambulance work (21.4%), administration (21.4%), and rescue work (10.5%).

Table 3-1 Demographics of respondents (N=1,492)

Demographic Variables	N	%

Gender		
Male	1413	94.7
Female	78	5.3
Age		
20 - 29	154	10.3
30 - 39	534	35.8
40 - 49	441	29.6
50 - 59	363	24.3
Final degree		
High school	354	23.7
Associate degree or 2-year college equivalent	400	26.8
BS equivalent	644	44.6
MS or equivalent	49	3.3
PhD or equivalent	2	0.1
Years of experience		
Less than 1 year	128	8.6
1 - 5 years	288	19.4
5 - 10 years	280	18.8
10 - 20 years	333	22.4
21 years or more	458	30.8
Specialty (Korea)		
Administration	319	21.4
Firefighting	696	46.7
Rescue work	156	10.5
Ambulance work	319	21.4

Instrument

The questionnaire used in this study was composed with an emphasis on the building life cycle. The questions that were used in frequency analysis later encompassed the entire building life cycles. The questions used in path analysis were only limited to the 'during construction' cycle. The following questions related to the building cycle were included in the survey: “How much does the building life cycle influence firefighting activities?”, “During construction stage only, how much does the completion degree influence firefighting activities?”, “What would be the building cycle in which firefighting is most challenging?”, “In which stage is the likelihood of fire occurrence highest? (During construction, During occupancy, or No occupancy for demolition)” For the other part of the survey, the respondents were asked to rate the importance level of factors that influence fire accidents on construction sites. The questionnaire was constructed around the building, fire, and human factors, which are the basic elements of a building fire, and there were detailed questions for each factor. A total of 15 items were used for the survey: 6 items for building, 6 items for fire, and 3 items for occupants (Table 3-2). Each item was scored on a five-point Likert scale from 1 to 5: “low important” (1), “slightly important” (2), “neutral” (3), “moderately important” (4), and “very important” (5).

Table 3-2 Survey items related to factors affecting fire on construction sites

Factors	Items	Alpha
Building	Available building information	.92
	The number of building entrances	
	Fire suppression features in and out of the building	
	Obstruction in the building	
	Falling through floor openings	
	Fire truck access to the building	
Fire	Securing sight	.94
	Falling objects	
	Explosion or collapse	

	Fire spread to nearby buildings	
	Fuel amount control	
	Fuel type control	
Human	The number of workers Worker's safety perceptions Cooperative actions of people at the scene	.85
Total		.97

Data analysis

Factor analysis

First, data screening procedures were executed and assumptions including the adequacy of sample size, factorability, presence of univariate and multivariate outliers, linearity, and multicollinearity, were checked. After the data screening procedures, a preliminary analysis of the underlying factors of the survey items was performed using principal component analysis (PCA). Since this method is the most widely used extraction method of component analysis and is most appropriate when the purpose is to reduce the number of items to a smaller number of representative components, the PCA method without rotation was used to determine the factor extraction (Costello and Osborne 2005; DeCoster 1998).

Path analysis

Path analysis (Wright 1960) is a type of multiple regression analysis technique that has been used to evaluate causal relationships between a dependent variable and independent variables in a model. By using this method, the researcher can estimate both the magnitude and significance of the causal relationship between variables. This analysis has the advantage of easily grasping direct and indirect effects that are difficult to grasp in regression analysis and focuses on the significance of each path rather than evaluating the suitability of the entire model. The results in a model can show a causal relationship by displaying the direct and indirect effects of independent variables on a dependent variable. Campbell and Cook (1979) explained that this method is appropriate when "theoretical, empirical, and commonsense knowledge of a problem" provides a mapping for the latent variables present and their probable causal links (Pajares and Miller 1994). Pajares and Miller (1994) suggested that path analysis is an appropriate approach when analyzing hypothesized relationships with both strong theoretical and empirical support. In the current study, the 'lavaan' package (Rosseel 2012) in R version 3.5.2 was used for the path analysis. This study focused on the significance of each path and the evaluation of the fit for the entire model. For selecting the appropriate model, Akaike information criterion (AIC) and Bayesian information criterion (BIC) criteria were applied, and smaller values of AIC and BIC indicated the model that had a better model fit with the data.

Structural Equation Modeling

Structural equation modeling (SEM) is a statistical method that has the advantage of being able to test multiple regression models at the same time as well as to test causal relationships of latent variables that are not observed, unlike path analysis, which deals with the causal relationship between observed variables. This technique combines factor analysis and multiple regression analysis to analyze the structural relationship between measured variables and latent constructs (Anderson and Gerbing 1988; Bentler and Chou 1987). Structural Equation Modeling analysis was performed to analyze the relationship between fire, building, and human factors after finding the most appropriate path among factors through the path analysis. The structural equation model comprises both the structural model (representing the causal relationship and correlation between latent variables) and the measurement model (representing the causal relationship between the latent variables and the observed variables). It is assumed that a latent variable consists of a component that can be explained by observed variables and the other component that cannot be explained. This study using R software with 'lavaan' package (Rosseel 2012) was performed for the structural analysis. In this study, the authors used the maximum likelihood method by assuming that the assumptions of univariate and multivariate normal distribution were met for three constructs: fire, building, and human. To evaluate model fit, fit indices, such as $\chi 2$, CFI, RMSEA, and SRMR, were comprehensively considered.

RESULTS

Preliminary analysis

The authors checked for missing data, univariate outliers, and multivariate outliers by using IBM SPSS 19.0. In this study, the method of deleting missing data (criteria: 80%) was chosen and 47 out of 1,654 cases were deleted. In addition, we found 62 cases as univariate outliers and 53 cases as multivariate outliers and deleted them. For identifying univariate outliers, z-scores of data were used and the data outside the observed between -3.3 to 3.3 were excluded. For multivariate outliers, Mahalanobis distance with a chi-square distribution was used. Therefore, the finalized sample size was 1,492 fire officers.

The reliability of the questionnaire

In this study, the authors used Cronbach's alpha method to assess the reliability of the survey items. Cronbach's alpha is a measure of internal consistency, which indicates how closely a group of items are closely related to each other. It is considered to be a measure of scale reliability. Typically, a Cronbach's alpha of .70 and above indicates good internal consistency, .80 and above being better, and .90 and above being the best. As shown in Table 3-2, Cronbach's Alpha values are calculated within the appropriate range. All items were calculated to be above .9 except for the human factor, which has a relatively small number of survey items. In particular, when all the items were calculated together, the most reliable was .97.

Frequency analysis

In this study, fire officers were asked about the effect of the building cycle on fire and fire suppression (Figure 3-1). Approximately, 82% of the respondents said that the building cycle had a significant effect on firefighting activities. In addition, more than 76% of the respondents said that the progress of construction has a great impact on firefighting activities. The answers to the questions about the most difficult times during the building cycle were: during the occupancy (57.1%), during construction (28.2%), and no occupancy for demolition (14.7%). When the respondents were asked in which building cycle is fire most likely to occur, approximately 73% indicated the occupancy stage, while only 20% indicated the construction stage

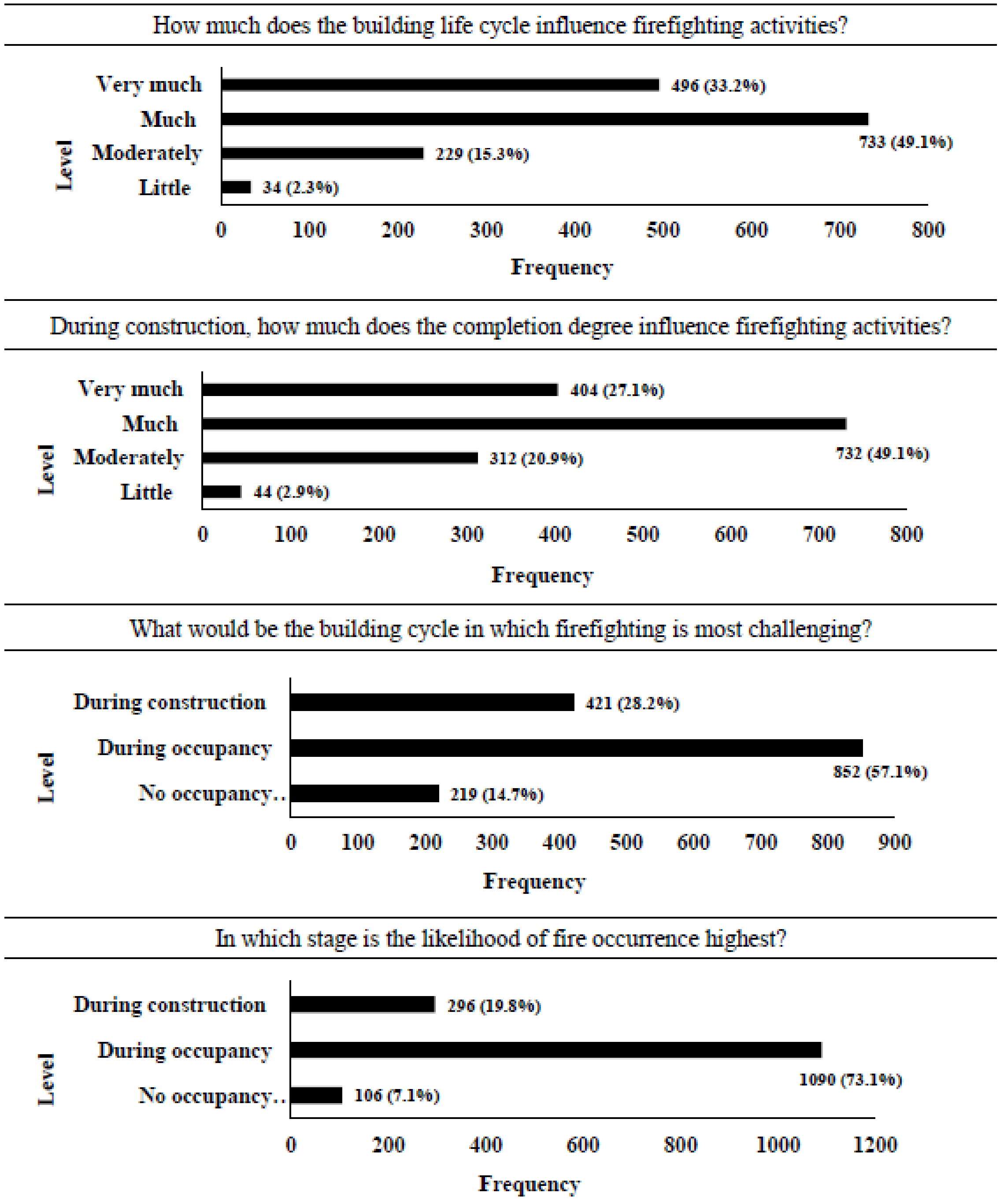

Figure 3-1 Questions related to the building cycle

This study analyzed how the perception of factors on construction fire safety differs with the working specialty of fire officers (administrative vs the field workers). Of the 1,492 respondents in this survey, there were 319 Administrative and 1,173 field workers. The authors used frequency analysis to report the proportions of the respondents who indicated a question was important. Of the five measures of the survey, ‘important’ and ‘very important’ were treated as important. The authors ranked the importance level for each factor based on the responses per respondent group. The ranking differences for each factor by working specialty are reported in Table 3-3.

Table 3-3 The difference in response by working specialty

Questionnaire	Administrative	Field	Diff.
Available building information	3rd	14th	11
The number of building entrances	4th	5th	1
Fire suppression features in and out of the building	13th	4th	9
Obstruction in the building	6th	8th	2
Falling through floor openings	15th	12th	3
Fire truck access to the building	2nd	2nd	0
Securing sight	9th	6th	3
Falling objects	11th	11th	0
Explosion or collapse	12th	13th	1
Fire spread to nearby buildings	14th	10th	4
Fuel amount control	8th	9th	1
Fuel type control	7th	7th	0
The number of workers	10th	15th	5
Worker’s safety perceptions	5th	3rd	2

Cooperative actions of people at the scene	1st	1st	0

Principal component analysis

The following process was used to verify the validity of questionnaire items. The Kaiser-Meyer-Olkin (KMO) and Bartlett's Test of Sphericity were used to evaluate the appropriateness of PCA. The value of KMO was .913, .902, and .767, respectively. It indicated that the data for building and fire factors were considered very good and the data for the human factor was considered adequate for structure detection (Pett et al. 2003). For all three factors, the value of KMO indicated that the data were acceptable for structure detection. In the case of Bartlett's test, it was statistically significant at α=.001 and it indicated that PCA may be useful for this data. After checking the assumption by using the KMO and Bartlett's test, various criteria were considered including the Kaiser Criterion, which states that factors with eigenvalues greater than or equal to one should be retained (Costello and Osborne 2005) and Cattell's Scree plot, which illustrates the factors and their corresponding eigenvalues. In this study, each factor had an eigenvalue larger than 1 (4.328, 4.660, and 2.331, respectively) and the result of the scree plot was the same. Additionally, the percentage of explained variance was investigated for each factor. If the percentage of cumulative variance in the first component or factor is 20% or more, the data can be described as represented by one component. Through this analysis, the building factor was 72.1%, the fire factor was 77.7%, and the human factor was 77.7%. Therefore, it was confirmed that each factor was well explained by the corresponding questions.

Path analysis

The results of this study presented six competitive models through path analysis. Three model groups were presented according to the position of the dependent variable among each factor, and there are two different indirect paths in each of the three model groups. For example, the dependent variable (building) for Models 1-1 and 1-2 (shown in Figure 3-2) are in the same position, but the indirect path has a different direction. The correlation coefficients affecting the variables in each competitive model are shown in Figure 3-2. First, the fire and human factors had direct and indirect effects on the building factor, respectively. In terms of the standardized coefficient values, the direct effects of fire and human on the building were 0.651 and 0.249, respectively, which were statistically significant. The indirect effect of the human factor on the building through the fire was 0.573 (model 1-1), and the indirect effect of the fire factor on the building through the human was 0.217 (model 1-2). Based on value, the coefficients of the direct effects of Models 1-1 and 1-2 are the same, and the indirect effects of Model 1-1 are relatively high. Second, the effects of building and human factors on fire were analyzed directly and indirectly. The direct impact of buildings and humans on fire was 0.643 and 0.351, which was statistically significant. Buildings and humans also have two different indirect paths that have an indirect effect on the fire. The indirect effect of the human factor on fire through the building factor was 0.529, and the indirect effect of the building factor on fire through the human factor was 0.315. Third, the fire and building sub-factors had direct and indirect effects on human factors, respectively. In terms of the standardized coefficient values, the direct effects of fire and building on humans were 0.541 and 0.379, respectively, which were statistically significant. Fire and Building also have two indirect paths that have an indirect effect on humans. The indirect effect of the fire factor on the human factor through the building factor was 0.329, and the indirect effect of the fire factor on the human factor through the fire factor was 0.519. The direct path coefficients of Models 3-1 and 3-2 are the same, and the indirect effect is relatively high for

Model 3-2. According to the information from model fit indices including χ^2 results, Akaike information criterion (AIC) and Bayesian information criterion (BIC) and theoretical backgrounds, the final model selection was conducted. Model 2-1 was chosen as the final model among competitive models for further statistical analysis of structural equation modeling since this model had the smallest values of AIC and BIC (AIC = 1105.978 and BIC = 1132.517).

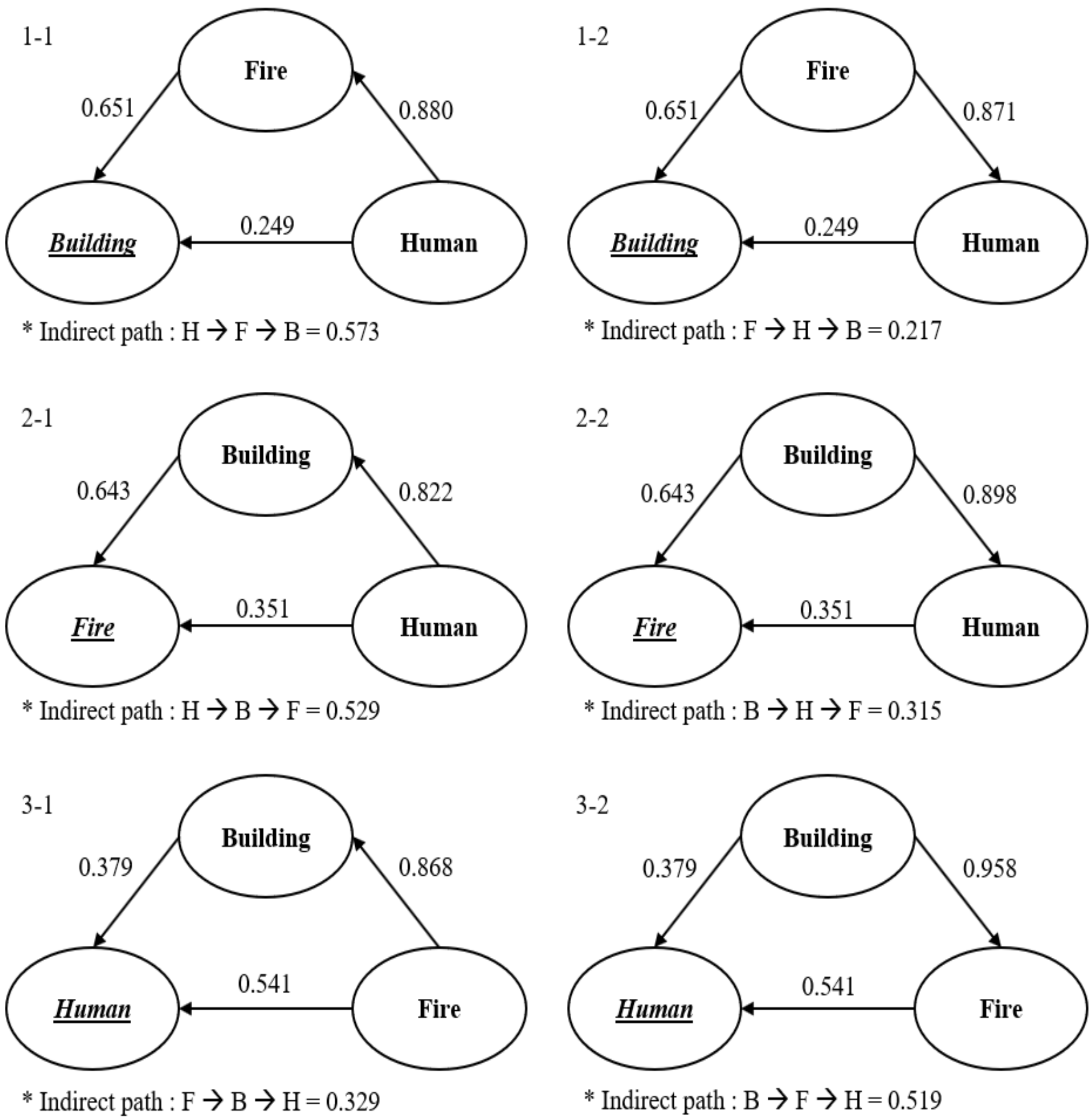

Figure 3-2 Path analysis competitive models

Structural Equation Modeling

Based on the relationship among factors through the path analysis, the measurement models of the latent variables were assessed. According to the initial investigation of results, adequate fit indices were calculated, and all values of factor loadings were statistically significant with $\alpha =$ .05 (χ^2= 2359.09; df = 74; CFI = 0.89; TLI = 0.87; SRMR = 0.058). According to the results from the measurement analysis, it can be concluded that all indicators adequately constitute the three latent variables: fire, building, and human. The 15 measurement indicators that form three latent variables were found to be statistically significant. Overall, it can be said that the analysis of the structural model is possible because the measurement indicators form the latent variables appropriately.

The structural equation model with the full sample of 1,492 fire officers yielded a good model fit (χ^2= 2359.09; df = 74; CFI = 0.89; TLI = 0.87; SRMR = 0.058). According to the results for each path (Table 3-4 & Figure 3-3), all coefficients were statistically significant.

Table 3-4 Standardized Coefficients from structural equation modelling

Path No.	Path	Standardized Coefficient*
(1)	Building → Fire	0.856
(2)	Human → Building	0.896
(3)	Human → Fire	0.151
(4)	Human → Building → Fire	0.767

*All coefficients were statistically significant ($p < .001$).

First, the building factor had a significant direct effect on the fire factor (β=0.856, $p < .001$) and the human factor had also statistically significant results (β=0.896, $p < .001$). Additionally, human factor had significant direct and indirect effects, respectively, on fire factor (β=0.151, $p < .001$; β=0.767, $p < .001$;). Based on the relationship among factors through the path analysis, it can be concluded that the indirect effect of the human factor on the fire through the building was relatively greater than the direct effect of the human on the fire. As a result of the verification of the structural model, all three paths were found to be statistically significant ($\alpha = .01$), so this model had high statistical significance. Therefore, overall, the structural analysis of this study is considered suitable.

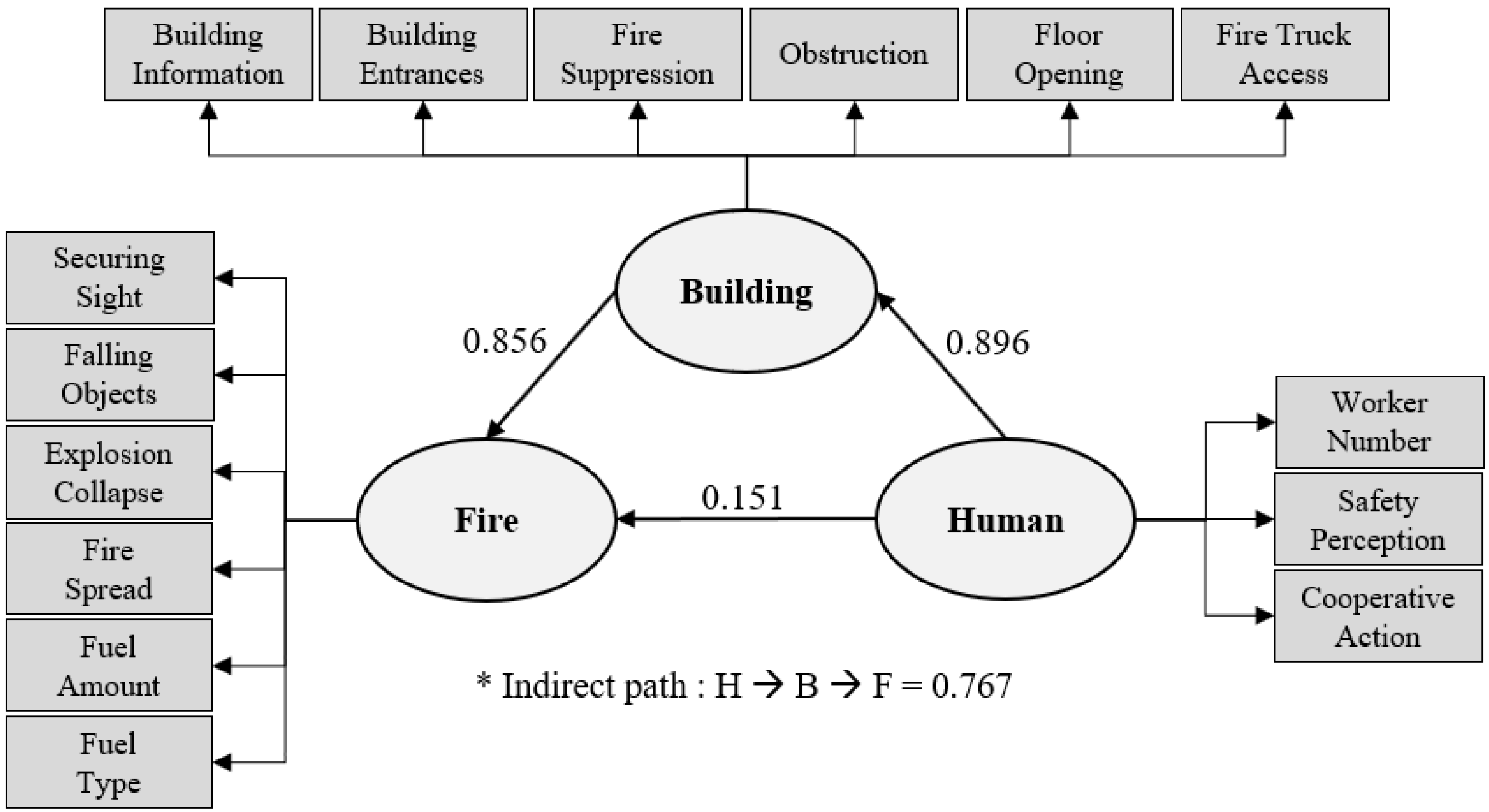

Figure 3-3 Structural Equation Modelling (SEM)

DISCUSSION

The impact of the building life cycle on fire

More than 82% of the fire officers surveyed said that the building cycle had a significant impact on fire officers' activities. However, most of the inspections for fire protection in buildings focus on the completion phase of the building. The regulations on safety planning of construction sites are not applied separately according to the progress of construction but are applied collectively regardless of the progress. Construction site workers may visit different times depending on their work, and various stakeholders such as suppliers, clients, and sub-contractors may visit and intervene in construction activities. In addition, construction activities and materials vary according to the progress of the construction site, which may cause a different risk of fire. However, it is difficult for them to know to what extent they are in the construction phase, and fire safety information and related training are lacking. This can make it difficult to implement an effective safety plan on the construction site.

Many fire safety regulations focus on 'During occupancy', but as the fire officers recognize, about 43% of respondents indicated that firefighting is the most challenging 'During construction' and 'No occupancy for demolition'. This finding indicates that fire safety regulations should be at least equally addressed during the construction and demolition stages. On a construction site where the fire safety system is not completely installed when construction workers perform activities in different locations, it is difficult to inform all workers of fire dangers. In order to efficiently alert workers to a fire accident on the site, an extra fire safety system must be installed in consideration of the characteristics of the construction site. In addition to the permanent fire and smoke alarms used during the building occupancy, portable alarms and efficient on-site monitoring systems can be viable alternatives. In addition, about 20% of respondents chose ‘During construction’ as the stage with the highest probability of fire. A possible reason for this is that construction sites usually contain various hazardous materials as well as activities that may cause a fire, such as welding, depending on the construction progress. Despite a slightly lower percentage compared to other stages, when it comes to fire safety, 20% still conveys a significant concern from the group. Therefore, the study of fire safety during construction should be a critical subject of study.

Main factors for fire on construction sites

The respondents answered that the most influential factors for fire safety on the site were 'cooperative actions of people at the scene', 'fire truck access to the building', and 'worker's safety perceptions'. The initial suppression is very important in any kind of fire accident, so the highest response rate was given to these factors. The 'cooperative actions of people at the scene' and 'worker's safety perceptions' items are included in the human factor. There is a limit to improving these two items included in the human factor only by modifying the building factor-based construction site fire safety regulations and behavior manual. Periodic and appropriate fire safety training on the construction site might be an effective approach. In the field, periodic fire safety training may be lacking due to financial reasons as well as schedule pressure, but the opportunity cost of the damage and impact of fire accidents caused by insufficient training can be much higher.

Also, workers' perception is a critical factor for fire safety as many studies have argued that 'worker's safety perceptions' is a very important factor in securing safety on the construction site. However, it is also an area that needs to be strengthened (Jafari et al. 2014; Kines et al. 2010; Mohamed et al. 2009). In the case of 'fire truck access to the building', this is expected to reflect the current characteristics of Korea. Most Korean cities have a high density of buildings, which makes it difficult for fire trucks to access the accident scene, which has a direct impact on the result of fire extinguishing.

Differences in perception of factors according to job specialty

When it comes to the perceived importance level for the identified factors affecting fire safety on construction sites, disparities exist among the respondent groups depending on the job specialty. In particular, two factors showed a big difference. First, in the case of 'Available building information', the fire department administrative personnel responded with it being the third most important factor among 15 factors, while among the field personnel, they rated the same factor as the fourteenth most important factor. Unlike in the past, a lot of building information such as BIM is required to be submitted to the fire department in charge before the start of construction. In response to this policy change, the administrative personnel considers such building information to be a very important factor in extinguishing the fire, but field personnel viewed it differently. If the changes in related policies are not accommodated in the field part, it will be difficult for advanced construction policies to contribute to the improvement of fire safety on construction sites. Continuous monitoring and follow-up on the effective application of the changed policy to the field part are required.

The second factor that showed a huge difference between the administrative and the field personnel was 'fire suppression features in and out of the building'. Field personnel perceived it as important because fire suppression features were always used at the fire site, but administrative personnel thought it was less important. The location of fire suppression features inside and outside the building is managed by building plans, and the fast and accurate provision of plans through the administrative part is a very important factor in suppressing fires. However, this perception disparity might also be attributed to the fact that building information including fire suppression features managed by the administrative personnel may not be efficiently transmitted to the field responders, and even if delivered, there may be information quality issues for the field personnel to properly use it. In Korea, administrative parts are in charge of related policies, but there may be limitations in effectively applying the policies in construction sites due to differences in perceptions between groups. Therefore, it is necessary to change related policies by considering field personnel's opinions and perceptions more actively to reassure their safety concerns.

Structural Equation Modeling

The authors analyzed the relationship among factors via path analysis and verified one specific relationship among comparative models based on the statistical results and theoretical backgrounds. The 2-1 model in Figure 2 was selected as the final model with the fire factor as the dependent variable and it indicated that the building and human factors were closely related to fire safety and firefighting activities on construction sites. Based on the results from path analysis, further analysis was conducted to confirm the relationship through structural equation modeling. The SEM model shown in figure 3 is the model proposed in this study. The building factor consists of building and fire design aspects and information related to the site and building.

Effect of Building Factors on Construction Site fires

Through the SEM of this study, it can be confirmed that design elements and building information can have a great influence on securing sight, falling objects, explosion or collapse, and fire spread when a fire occurs on the construction site. For example, the fire suppression feature in and out of the building factor is the item most closely related to the spread of fire on the construction site, and when properly controlled, it can effectively prevent the spread of fire on the construction site. In addition, securing sight in the event of a fire accident is one of the most important factors for evacuation and extinguishing. Building entrances and obstacles are elements that can affect the visibility of evacuees and firefighters. In addition to this, building design for fire safety can improve visibility in case of fire. For sufficient securing sight during the evacuation from a building, a design that slopes the ceiling to contain smoke for a certain period can be an effective alternative method.

Many materials that can become falling objects in the event of a fire on the construction site are stored in various places on the site. However, since it is difficult to check the updated information on the storage location of construction materials according to the progress of the construction project, the risk of falling objects on the construction site can be an obstacle to fire suppression and evacuation. In order to effectively manage the risk of falling objects on the site, it can be an efficient alternative to select a space to store construction materials at the building design stage and provide this information to the site and the fire department. This alternative can also prepare for the risk of collapse due to an explosion, as it informs in advance the location of materials that could cause an explosion in the event of a fire on a construction site. The fire truck access is an essential part of extinguishing fires on the construction site in the early stages, which is closely related to fire spread.

Also, “the available building information” could be closely related to the basic information for “fuel type control” and “fuel amount control” classified in fire factors, which could be determining factors in controlling explosions on construction sites. The available building information includes advanced data such as BIM and 4D BIM. Additional building information such as BIM should be used in the safety aspects of the construction site beyond data for design quality verification and building permits. Since 4D BIM data includes scheduling of construction sites, fire risk information for each construction period can be provided to the site. In the fire department, advanced building data such as BIM should be used for 3D map-based mock training and simulation. To effectively use this advanced information to improve fire safety on construction sites, the information must include factors that can affect the fire, such as fuel type and amount. BIM guidelines for fire safety management of construction sites beyond the traditional use of BIM should be developed.

Effect of Human Factors on Construction Site fires

On the other hand, in the case of the human factor analyzed in this study, rather than directly affecting the fire on the construction site, it was found that it indirectly affected the fire on the construction site through the building factor. For example, improved worker's safety perception has an effect on the increased awareness of building factors, such as building information, fire suppression features, entrances, obstruction, and floor openings of a construction site, that can be proactively addressed during design and design review stages to control fire risk. The cooperative action of construction workers can be an important factor for the fire truck to quickly access the site. In addition, the number of workers on a construction site can affect the location and number of entrances to the building, which is necessary information for efficient evacuation in case of fire.

CONCLUSION AND RECOMMENDATIONS

Most building fire regulations and codes are established for buildings under the occupancy phase and the construction stage is inadequately considered. Previous studies that analyzed fire safety on completed buildings have mainly focused on the effectiveness of fire safety systems and evacuation plans. However, these studies have not yet explored the factors that influence the fire risk of the construction site or how they can improve fire safety. In addition, there is a limitation that survey-based studies related to the safety of construction sites contain only the opinions and perceptions of construction workers. In addition to exploring the level of importance of building life cycle in relation to fire safety, this study examines the importance of various building, fire, and human factors in determining fire incident results as well as the relationships between factors through a questionnaire. Specifically, research was conducted focusing on fire officers who are closely related to fire accidents on construction sites, and this is the first study to investigate fire officers' perceptions regarding fire safety on construction sites. This study provides some information on the relationship between the importance of fire safety and the factors influencing fire in the construction stage setting not previously studied. This study contributes to the body of knowledge by providing information on the importance of factors affecting fire on construction sites and the relationship between factors.

The results of the study suggest that fire safety on the construction stage is as important as a completed building, and in particular, field firefighters answered that an effective fire safety system on the construction site is important. There were also some noticeable differences in the perceived importance of factors affecting fire safety on building construction sites depending on the fire officer's job specialty. Through the SEM model, this study found that building-related factors had the greatest direct effect on fire-related factors concerning fire safety on construction sites. Although human factors did not have a strong direct relationship with fire factors, the human factors had an indirect impact on fire factors through the human factor as a mediator to achieve fire safety. The SEM model proposed in this study can be used as a theoretical framework for future related studies. These results provide clear evidence that the field of fire safety on the construction stage should also be studied and developed with a similar interest to the field of fire safety in completed buildings. By knowing the effect and relationship between the factors contributing to fire safety on construction sites, more effective fire safety measures, such as policies and regulations, can be developed.

The methodology and interpretation chosen for this study presented some limitations for analyzing the results. The current study on fire safety focuses on the perception of fire officers under the context of building construction sites located in densely populated areas. Future studies should account for other stakeholders involved in the fire safety realm since the perception might be different depending on the occupational group. Additional research may be required for clients, designers, sub-contractors, building regulators, and equipment suppliers. The SEM model can be either modified or further validated by collecting data from various occupational groups.

Moreover, the density of surrounding buildings, type of construction, climate, and the number of onsite workers can vary with region, which may lead to a difference in survey participants' perception. Applying more powerful data collection and sampling techniques to provide a larger sample size reinforces the results of this paper. Therefore, future studies can take the above-mentioned characteristics associated with the construction sites into consideration so that construction site fire safety can be addressed on a larger scale.

CHAPTER IV

DEVELOPING DATA REQUIREMENTS OF BIM FOR FIRE SAFETY ON CONSTRUCTION SITES

ABSTRACT

Fire risk on construction sites is a threat to property and construction workers' safety. To minimize the loss of human life and property, effective fire safety plans should be established on the construction site. Despite the increased use of building information modeling (BIM) in the architecture, engineering, and construction (AEC) industry, using BIM to improve fire safety during the construction stage is lacking. The uses of BIM for construction safety are limited to the role of a rule checker, such as detecting holes in a plan to identify possible safety hazards. In addition, there are limitations in studying fire and human evacuation performance by reproducing actual fire accidents. Therefore, computer simulation can be a cost-effective approach for calculating evacuation performance and fire growth on a construction site. BIM-based simulation studies for improving fire safety on construction sites have not been sufficiently studied. Using BIM data to improve construction site safety, such as fire safety, is one of the most effective methods. This study aims to provide a BIM-based simulation framework to simulate fire and evacuation performance on construction sites. To propose a framework, the author analyzes how to apply BIM in simulations related to fire safety on construction sites and classifies BIM-based input data required for simulation. This study contributes to the overall body of knowledge by providing a BIM-based simulation framework that can efficiently manage fire safety on construction sites.

INTRODUCTION

Due to various economic reasons and advances in construction technology, complex and taller buildings are being built in densely populated urban areas. The advancement in technology has the advantage of meeting the residents' needs of residents by providing a variety of building designs as well as newer systems in the building that improve the life quality of occupancy, but it also comes with the disadvantages of complicating the project. As the construction project has become complex, it has become difficult to track the design of the project, which can lead to various problems such as violation of regulations and design errors. BIM data is being used as a solution that can proactively consider these issues in the design stage (Koo and O'Connor 2022). New regulations have been established to mandate the use of BIM from administrations such as Korean governments for maintenance and energy management of buildings and it is expected that this type of mandate will gradually increase considering the advantages of BIM (Rehman et al.). In addition, as the number of large construction sites increases as well as the construction activities and materials stored on the site diversify, various accidents such as fire accidents are occurring (Kim et al. 2019). Fires on building construction sites, different from other types of building fires, have many unique characteristics, such as limited movement space and uncertainty of exits (Kim et al. 2021). The Bureau of Labor Statistics (BLS) recently released the Census of Fatal Occupational Injuries (CFOI), and the construction industry topped the list with 1,008 worker deaths (Kim et al. 2021). In the National Fire Protection Association (NFPA) report, the frequency of fires to construction among all building fires is reported to be about 1%, but direct property damage is reported to be about 2% (Kim et al. 2021). This shows that accidents on construction sites cause more financial damage than the frequency could indicate.

Building information modeling (BIM) is a data process developed by various technologies and tools involving the generation and management of digital data of the physical and functional characteristics of spaces (Volk et al. 2014). Research on the use of BIM data for construction sites tends to focus on the quality of the design. In a study, the authors developed a BIM-based quality-checking system for each stage of construction to improve the quality of architectural design and applied it to the entire construction process (Choi et al. 2020). In order to efficiently manage design quality in prefabricated modular construction projects, there is also a study that suggested a method of managing the project by matching the 3D BIM model with the actual construction site (Li et al. 2020). Furthermore, research on the use of BIM data has been predominately limited to building energy analysis and rule checking (Carvalho et al. 2021). However, BIM data has the flexibility to expand, and when it is used efficiently, it can help to increase safety on the construction site (Lu et al. 2021; Tran et al. 2022). As safety on construction sites is directly related to both life and building property losses, it is expected that BIM implementation by building designers will increase due to BIM's capability to improve safety through better design review and preconstruction coordination (Teo et al. 2016). Currently, research that analyzes accidents on construction sites using BIM has been conducted to detect design errors and eliminate risks (Shen et al. 2022). These studies utilize BIM data but have limited application to verify design plans and codes. There is a study that suggested the development of system information for building code checking at the permitting phase to improve building code compliance using BIM (Yogana and Latief 2021). There is also a study on a system that checks whether the building code is affected when a part of the design is changed during the construction phase (Kim et al. 2019). One of the reasons attributed to the limited utilization of BIM data for such purposes is that designers do not have clear data definitions and guidelines for generating BIM data related to safety on construction sites. In order to use BIM data to enhance the safety of fires and evacuation of construction sites, it is very important to define BIM data for site safety and provide it to designers so that the porper data can be created or entered during the design phase.

An appropriate fire safety plan for a construction site is key to preventing fire accidents and increasing evacuation performance in case of fire. However, it is not realistic to study fire safety performance during a real fire on a construction site, and there are many limitations. Therefore, computational simulations are widely used as an approach to assessing fire risk and human evacuation on construction sites. Virtual scenario-based fire and evacuation simulations provide a low-cost and low-risk approach to predict evacuation performance to enable safe design and emergency management on construction sites. Fire and evacuation simulation can be applied at all stages of construction, from the design stage of the project to the operation and maintenance stage after the building is completed. The simulation outputs can also be used to evaluate fire safety performance during the construction stages to support evacuation planning on construction sites. However, fire safety simulations on construction sites have limitations in not sufficiently reflecting interoperability with BIM. There are several obstacles, including the interoperability of BIM data and the technical limitations of simulation software currently available. This study aims to suggest a BIM-based integrated simulation framework that can implement fire and evacuation simulations to improve fire safety on construction sites. Therefore, in this study, data to be included in BIM during the design was presented in order to enhance fire safety on construction sites. In addition, in order to effectively use BIM to improve fire safety on construction sites, the author has provided BIM data related to fire safety on site. The BIM data structure was analyzed, and the data necessary for the analysis of fire accidents and workers' evacuation were reflected in the guidelines. By providing BIM data for construction site fire safety to stakeholders including designers during the construction project planning stage can help create and utilize BIM data for analyzing fire and evacuation safety on construction sites.

BACKGROUND

Construction accidents

The construction industry is known as one of the most dangerous industries when it comes to injuries and fatalities. According to the 2019 Occupational Safety and Health Administration report, 20% of all industrial worker deaths are reported in the construction industry while construction accounts for 5-6% of the US GDP (BLS 2019). The construction industry results in fatal and non-fatal injuries from a variety of activities, which cost construction companies millions of dollars per year (Alkaissy et al. 2022). The leading causes of working fatalities in the construction industry are falls, being-struck-by-an-object, electrocution, and being-caught-in-objects, a.k.a the “Fatal Four.” These accident types are responsible for 59 percent of all construction worker deaths (OSHA 2020). According to the Centers for Disease Control and Prevention of all industries, construction causes the most fatal fall accidents, accounting for 51% of all fall accidents in the United States (OSHA 2020). Also, one fatal injury at the construction site costs an average of $991,027 in hospital costs (Sidani et al. 2021).

Many studies to accidents on construction sites analyze causes by accident type. Using statistical data from the Bureau of Labor Statistics (BLS), researchers investigated the types of fatal construction accidents and found workers to die mainly from falls (Zhou et al. 2008). Hallowell and Gambatese (2009) used the Delphi process to identify key safety-critical activities in formwork construction work. According to this study, the most dangerous activity was exposure to harmful materials. In addition, several approaches have been proposed by introducing the concept of evaluating and assessing safety on the construction site. For example, Yang, Chew et al. (2012) analyzed all the construction accidents in the U.S. from 1995 to 2008 and designed a system to assess the accident possibility. They provided the designed system and confirmed that some construction site accidents can be predicted using past statistical data. If workers are aware of the safety or risk issues on the construction site in advance, workers tend to behave safely (Mohamed et al. 2009). Thus, predicting construction site accidents and informing workers through training is a very important points to improve the safety of construction sites (Abdullah and Wern 2011; Cooper and Phillips 2004; Hallowell and Gambatese 2009). As such, in order to effectively alert construction workers to the risks on the site, it is important to analyze the risks on the construction site. One of the ways to improve safety on a construction site is to use safety simulations to predict and prevent accidents. Providing a safety assessment model based on fire and evacuation simulations can significantly improve the safety of the site. It can also provide economic benefits to the construction industry. The construction industry is inherently dangerous, but the prediction of risks on the site using approporiate data can mitigate some of these risks.

Fire accident on construction sites

Although fire accidents on construction sites are less frequent than the main types of accidents such as falls and collapses, their importance should not be overlooked (Su et al. 2021). A fire accident on a construction site can cause secondary accidents such as collapse, which can increase human and property damage. Fire safety studies revealed that unsafe management of construction sites is a major cause of fires (Wang et al. 2021). Also, hazardous activities on construction sites, such as welding, can contribute to a fire (Doglione and Piccinini 2022). Selecting the proper location and method of storing combustible materials on construction sites was considered to be an important factor in preventing fire accidents (Sielicka et al.). In this study, it was pointed out that improper site layout and installation of heating devices that did not consider the storage location of combustible materials were the main causes of fires on construction sites.

Studies have attempted to approach methods for reducing fire accidents on construction sites in terms of construction worker behavior. In one study, a survey of 215 construction workers in New Zealand was analyzed and a construction worker safety behavior model was developed and tested (Guo et al. 2016). The results showed that production pressure was a critical factor that has direct effects on safety motivation, safety knowledge, and safety compliance including fire safety. And one study provided a way to predict the fire risk of a construction site using ten safety conditions determined through a literature review (Patel and Jha 2015). As a result, safety climate constructs such as manager supervision, work pressure, employee engagement, and awareness of fire risk were significant relationships with worker safety behavior.

Recently, fire safety plans for construction sites are presented based on various technologies and theories in other fields. For example, a study proposed a model to evaluate the fire risks on construction sites using artificial intelligence after collecting data at a construction site using the Delphi method (Ayhan and Tokdemir 2019). The model was able to predict 84% of the accident results on construction sites. In addition, it predicts the safety status of construction sites by using information measured on site in real time rather than existing statistical data. One example is the study of a construction site using a Real-Time Location System (RTLS) to develop a system that proactively detects and alerts workers to fire hazards on the site (Li et al. 2016).

Simulations related to fire safety

There are two main types of simulations used to evaluate fire safety performance on construction sites. One is fire simulation, a fire dynamics modeling tool used to calculate fire fluid flow (Kallada Janardhan and Hostikka 2019). Recently, many studies have shown attempts to utilize BIM data for fire simulation (Dimyadi et al. 2018; Sun et al. 2022). BIM files can be utilized to visualize the results in 3D within the fire simulation (Rahmani and Salem 2020). Fire simulations have been used for a variety of purposes, such as assessing the fire safety of existing buildings, and analyzing factors that influence the growth of building fires (Ahn et al. 2019).

Another type is evacuation simulation. Evacuation simulation can design an evacuation scenario using the interaction of agents and space and can also reflect human behavior (Cheng et al. 2018; Folk et al. 2019; Şahin et al. 2019). The algorithm of evacuation simulation can simulate the behavior of occupants through mathematical (Haghani 2020), theoretical (Ibrahim et al. 2019), and logical methods (Yang et al. 2020). This application has been applied to various disaster situations such as tsunamis, hurricanes, and earthquakes as well as the evacuation of buildings (Feng and Lin 2022; He 2021; Wang and Jia 2021). Many studies have been conducted on evacuation simulations based on various disaster scenarios. For high-rise buildings, which have a longer evacuation completion time compared to other building types, evacuation and fire simulation studies have been actively conducted. In high-rise buildings, most of the occupants use the stairs in a disaster situation, so there is a high probability that a bottleneck will occur near the stairs (Gerges et al. 2021). In order to alleviate the bottleneck, a study suggested the evacuation sequence of high-rise buildings and the management of evacuation passages without obstacles

(Zang et al. 2021). In addition, evacuation simulation is used by assuming an emergency in public places such as subway stations. Underground facilities such as subway stations have a fatal weakness for evacuation because smoke flows in the same direction as people evacuate (Chen et al. 2021). In many studies, fire and evacuation simulations were performed assuming various emergencies in subway stations (Zhang et al. 2019). The results of this study suggested that the width of the exit did not significantly affect evacuation in subway stations, unlike general buildings, and that the number of people on the platform and the number of passengers on the train should be limited to a certain number (Qin et al. 2020). Recently, attempts have been made to reflect on factors that influence human behavior in evacuation simulations. Many studies have been conducted to graft theories of psychology and sociology to the agent model used for simulation (Li and Guo 2021; Şahin et al. 2019). Nevertheless, their research has implied difficulties in quantifying human responses to disasters and applying them to simulations to predict fire safety.

The main structure of the evacuation simulation considering the fire in the building can be classified into fource categories: a space/structure model that reflects the space in a building, a movement model that reflects information about occupants, a behavior model that reflects the behavior of occupants by decision-making algorithms, and a fire data model that reflects fire analysis. Table 4-1 shows the main structures of the fire evacuation simulation in a building.

Table 4-1. The main model structures of the building fire evacuation simulation

Type	Definition	Detail
Space/Structure Model	• By dividing the space of a building, the structure of each space is created. Through this, it is possible to determine whether the occupant occupies the space.	• Space network • Fine network • Continuous space

Movement Model	• A model of the movement of occupants that is affected by the calculation of the simulation algorithm	• Density • User's choice • Distance • Unimpeded flow • Cellular automata
Behavior Model	• It is a behavioral model of the occupant and can make appropriate changes depending on the evacuation speed and situation.	• Probabilistic • Artificial intelligence • Conditional • Behavior • No behavior
Fire Model	• It is an analysis of the fire performance of a building, and it is divided into four types by the method of reflecting the fire data.	• Cannot incorporate fire data • Can import fire data from another model • Allow the user to input specific fire data • Use own simultaneous fire model

BIM based construction research

BIM has recently become a new paradigm throughout the construction industry. Because of its capability in data integration, BIM has been used for building information management and performance analysis IFC is an international standard for BIM data and is an important data structure for representing BIM. Extracting necessary data elements according to the purpose of building performance analysis in IFC is essential for the efficient use of BIM (Carbonari et al. 2018). In addition, when BIM is used in construction projects, the same information source can be delivered to each sub-step of construction, improving inefficiencies in information management such as manual data re-entry. BIM makes it possible to standardize spatial data information exchange in all fields to construction, and ultimately, through feedback with basic design data, design quality and management quality can be improved (Ozturk 2018). This enables technical collaboration between various work fields required throughout the life cycle of a building and provides a technological basis for changing the collaboration method that has relied on the experience of designers in the relevant field to the establishment of an information-oriented collaborative method process (Ali et al. 2022). Each country developed its BIM guideline considering local characteristics, but fire safety was excluded in the BIM guideline development. Table 4-2 shows the BIM guidelines developed and operated in each country.

Table 4-2 Current status of BIM guidelines

Country	Guideline Title	Development agency
USA	National Building Information Modeling Standard	NIBS
	BIM Guide Series	GSA
	General Building Information Handover Guide	NIST
South Korea	Facility Business BIM Basic Guidelines	Public Procurement Service
	BIM Guidelines for Architecture	Ministry of Land, Transport and Maritime Affairs
	Construction Information Model Common Criteria	KICT
Finland	BIM Requirement 2007	Senate Properties
Denmark	Digital Construction	NAEC
Germany	BIM/IFC User Guide	Germany IAI
Norway	Information Delivery Manual	Statsbygg
Singapore	Core Net	Singapore construction agency

The fire safety in the construction industry is a field in which the application of construction information such as BIM is relatively slow compared to other fields. Recently, as the number of buildings receiving safety certifications in various countries increased, the quality improvement of buildings is becoming the goal of building design (Ku 2019; Mason et al. 2019).

In addition, as building safety regulations are strengthened, building design may be governed by related regulations. In the existing building performance analysis process, the user modifies the design drawing to fit each performance analysis simulation and then performs the simulation. However, in the BIM-based process, the modeling information itself can be analyzed and transformed, so the part of the designing simulation, which has been the domain of experts, can be changed more openly (Pruskova and Kaiser 2019). Since the introduction of BIM, there have been studies trying to develop an international standard for building performance evaluation and fire simulation based on IFC (Farghaly et al. 2022; Shi et al. 2019). However, since IFC aims to contain information from all areas of a building, it may be inefficient for use in certain areas such as fire safety or evacuation. Currently, in the case of simulations for building fire safety, there are many limitations in interworking with BIM. The range of interworking is narrow enough to extract and analyze only some 2D information from the BIM file or visualize the analysis results in 3D.

BIM data in simulations related to fire safety

Most of the existing fire and evacuation simulation programs are based on two-dimensional plans, and the user interface is not intuitive, so it is not possible to provide an environment that can use BIM data more efficiently for simulations (Sun and Turkan 2020). For accurate fire simulation, it is better to share and use the same spatial information based on BIM spatial information between evacuation and fire simulation at the same time. (Chen et al. 2021; Ma and Wu 2020; Sun and Turkan 2019). In the case of using two-dimensional spatial data, additional data cleaning work is required, which is convenient for users. However, in order to utilize BIM data in simulations, it is inconvenient that the user has to create BIM spatial information from 2D spatial information (Shams Abadi et al. 2021). In the fire and evacuation simulation, the spatial information of the real space must be accurately implemented in the simulator, and the sure way to check this accuracy is for the user to visually compare the real space with the spatial information implemented in the simulator (Huang et al. 2019). Therefore, a simulator using BIM data, which is based on 3D spatial information, is essential not only for user convenience but also for improving the accuracy of the simulation.

Also, in terms of fire safety on construction sites, adaptability to the future architectural design environment through BIM is necessary (Kincelova et al. 2020). BIM can check all information related to 3D modeling, 2D plans, materials, and construction at once, which can help with design and construction changes (Gaur and Tawalare 2022). Considering the advantages of BIM such as ease of use, work efficiency, and wide application, many aspects such as fire safety and evacuation will be actively integrated into BIM. In addition, existing fire safety simulation tools will be adapted to the integration with BIM. Therefore, it is essential to study which variables should be included in BIM to improve fire safety on construction sites. Table 4-3 summarizes the functions of the fire evacuation simulations currently developed.

Table 4-3 Functions of fire evacuation simulations software

	Simulex	Building EXODUS	FDS+EVAC	STEPS	PATHFINDER
Modeling Method	Partical Behavioral Model	Behavioral Model	Movement Model	Partical Behavioral Model	Partical Behavioral Model
Grid/ Structure	Continuous Space	Fine Network	Continuous Space	Fine Network	Continuous Space
Import Geometry	CAD	CAD	N/A	CAD/BIM/ FBX/ Sketchup	CAD/FDS/ Pyrosim

Visual Simulation	2D	2D/3D	2D/3D	2D/3D	2D/3D
People Behavior	Implicit	Conditional	Conditional	Conditional	Conditional
Escape Route Guide	No	Yes	No	No	No
Toxicity	No	Yes	No	No	No
Visibility	No	Yes	No	No	No

Point of departure

Currently, the use of BIM in the construction industry is rising, and BIM has been applied to various simulations. BIM technology can create data from all the spaces contained in a building, and based on this, advanced results such as 3D drawings can be derived. In addition, it can be used for various simulations such as the amount of material required, the cost, the construction process, and energy analysis. BIM research is steadily progressing in the construction field, and the use of BIM is expected to gradually increase due to the efficiency of the use and exchange of building data. However, studies related to BIM to improve fire safety on construction sites are lacking. Until recently, the existing CAD file or image file was mainly used for spatial data for fire and evacuation simulations. As the use of BIM in the construction industry increases, the linkage between fire safety simulation and BIM is also important, but there are few simulation software that can seamlessly link with BIM.

In this study, BIM data required for fire and evacuation simulations of construction sites are identified. This study identifies variables that can be provided in BIM among various variables included in input data for fire and evacuation simulations. This study contributes to the overall body of knowledge by providing the classification of BIM data necessary for fire and evacuation simulations to improve fire safety. The classification of BIM data required for fire and evacuation simulations can be a basis for improving fire safety on construction sites. Therefore, this study facilitates the use of various relevant simulations for fire safety management on construction sites, ultimately helping to minimize or prevent human life and property loss from building fires.

METHODOLOGY

BIM-based fire and evacuation framework

The goal of this study is to propose a method to integrate fire and evacuation simulations and develop a BIM-based integrated simulation framework to improve fire safety management of construction sites. Figure 4-3 is the overall framework provided in this study. This prevents fire accidents on construction sites and helps to reduce human and property losses in case of fire. In the first step of the framework proposed in this study, data on the shape of the building are required. In the case of IFC, the standard model of BIM data, it includes all data throughout the building, including architectural design, construction, and structural facilities. The IFC standard was developed to express various data generated throughout the construction of a building. IFC is a model that not only deals with physical building elements such as walls, doors, and windows, but also electronically expresses various information such as schedules, space elements, and costs. IFC is divided into four layers (Domain, Interoperability, Core, and Resource) that express different levels. Therefore, a lot of data reduction, simplification, translation, and interpretation is required to make the IFC model compatible with a specific performance analysis model such as fire safety. A specific data format developed for compatibility with specific building performance analysis programs will be required. Figure 4-1 shows the data compatibility process between the design program and the analysis program using BIM data. The specific format in Figure 4-1 describes the data format required for the fire and safety simulation of the construction site.

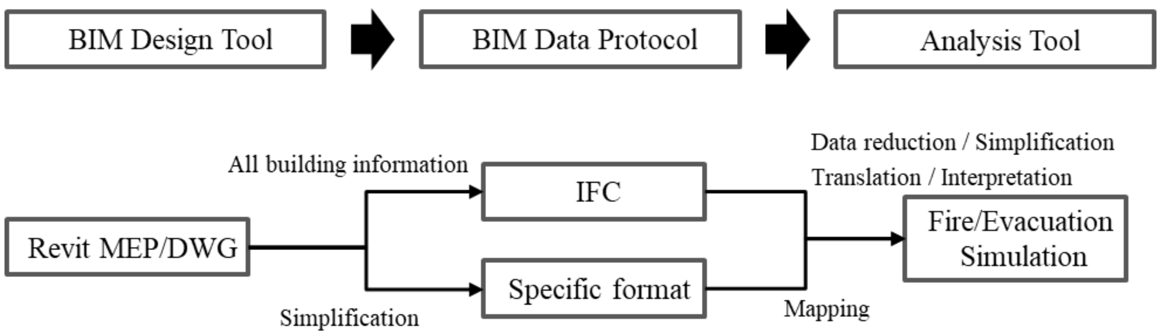

Figure 4-1 Data compatibility process between BIM software program and analysis program

In the case of BIM-based building fire and evacuation simulation, the simulation should be performed based on the integrated model created in BIM software. To this, it is necessary to select a building fire and evacuation analysis tool and ensure interoperability with BIM software based on IFC or a separate standard format. However, until now, the standard format and complete compatibility of fire and evacuation simulation tools are not possible, so it is necessary to define the range of information to be converted for fire performance analysis in the BIM model. Based on this, it is necessary to define the information form of IFC or specific format and investigate whether the file created from the BIM software is suitable for the fire and evacuation performance evaluation of the building. The 'IfcRelDefinesByProperties' is the IFC object representing the building spatial information and property information required for fire and evacuation analysis in this study. This object contains IfcBuildingElement and IfcPropertySet as sub-objects. IfcBuildingElement represents the structure in the building, and IfcPropertySet represents each property information about the building object. That is, through this object relationship, the property information of each structure can be defined from IFC. Figure 4-2 shows the relationship of IfcRelDefinesByProperties through the IFC browser in the form of a tree menu, and material properties can be defined in a specific structure through IfcPropertySingleValue.

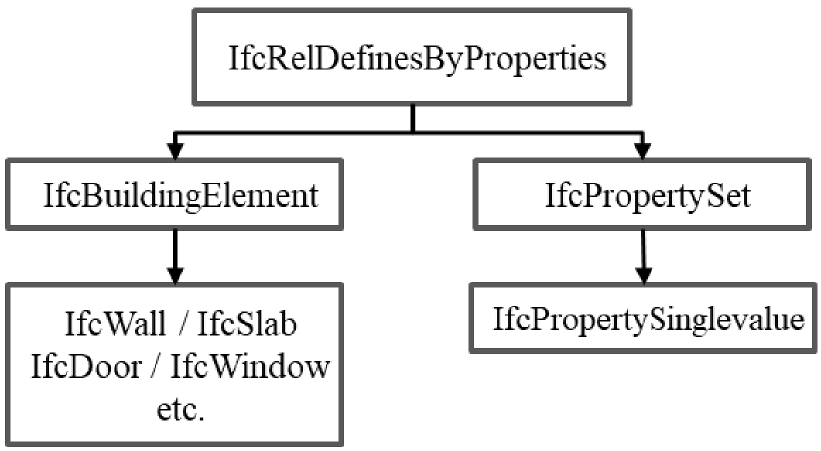

Figure 4-2 Tree relationship in IfcRelDefinesByProperties

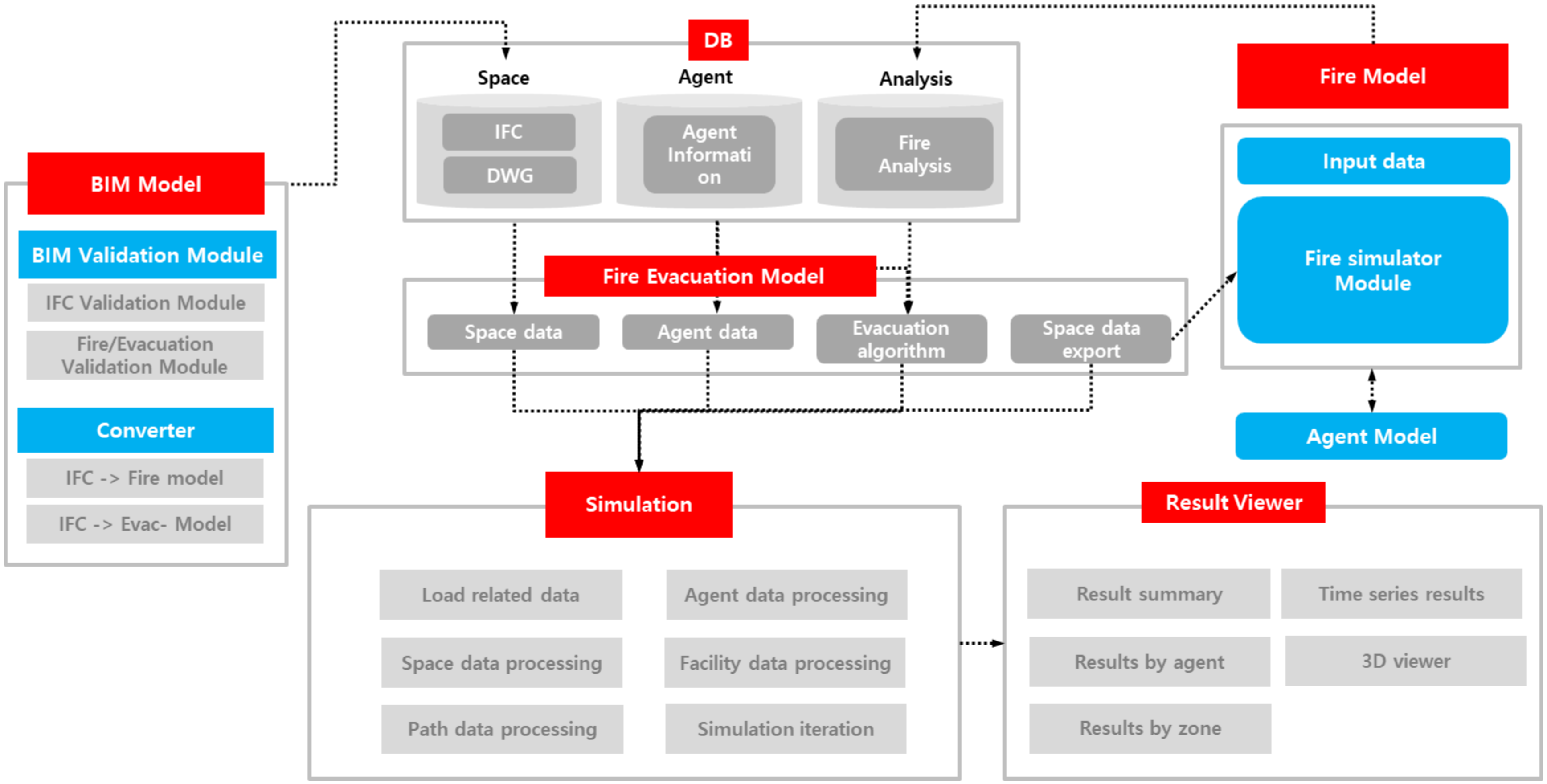

Figure 4-3 BIM-based fire and evacuation simulation framework

Classification of BIM data to improve fire safety on construction sites

In this study, among the input data required for fire and evacuation simulations, variables to be considered in the BIM stage were analyzed. A typical application of BIM data to improve the fire safety of buildings is to utilize the data for fire and evacuation simulations. Although the main purpose of BIM is to create and manage data during the design, construction, and operation phase through improved visualization of projects, BIM data has wide extensibility that can be directly used for various simulations. In order to fully utilize BIM data, it is essential to provide designers with the data requirement of BIM for fire safety. Therefore, in this study, the BIM data to be established for the efficient use of fire safety simulations for construction sites was classified, and detailed data necessary for each class was provided. For this purpose, the variables that can be considered in the BIM stage among the main input data required to perform the algorithm of building fire and evacuation simulation are classified into 5 classes. This study provided a conceptual map of the space of a building, which is mostly focused on the BIM stage. In addition, categories, variables, data types, and descriptions included in each class are provided intuitively.

RESULTS AND DISCUSSION

Utilization of BIM data to improve fire safety

In this study, in order to efficiently use BIM data for improving fire safety on construction sites, the current state of fire and evacuation simulations was analyzed. Table 4-4 shows the generational changes and main features of the evacuation simulations. It is currently classified as the 4th generation model, and this stage provides linkage between fire and evacuation simulation. In this stage, various data such as BIM, human behavior, and fire dynamics are applied to fire and evacuation simulations to obtain a more accurate analysis. However, the analysis of what data should be extracted for fire safety simulation from BIM generated during the building design stage is insufficient. For efficient use of BIM data for fire safety, it is necessary to classify BIM data that affects the operation of fire and evacuation simulations and provide them to users.

Table 4-4 Generation and characteristics of fire and evacuation simulation

Generation	2nd Generation Model	3rd Generation Model	4th Generation Model
Main Features	• Computer-based hydraulic model • Consider complex conditions such as queues • Conditions inside the space are not clearly reflected • In the case of evacuation simulation, it is impossible to consider individual characteristics	• Describe the flow of the population using liquids or balls (Ball bearing model) • Influenced behaviors assumed by the physical environment • Lack of decision-related factors	• Various behavioral variables can be considered • Behavior is influenced by physical, psychological and social conditions • Linkage between fire and evacuation simulation • Use of advanced data such as BIM

In this study, variables that can be used in fire and evacuation simulations among BIM data representing spatial information of buildings were investigated. Not only the information included in the algorithm of the simulation but also the internal classes and data structures that can exchange data with other simulations are shown. Most simulators accept input data from the main algorithm and perform simulations based on them. Also, when receiving external data, the external data is converted to the input data format of the main body and then transferred to the main body. The main class used inside the fire and evacuation simulator can be divided into five categories.

- Space class: This is a class module that manages the physical information of the building space.

- Path class: This is a class module that manages path information.
- Facility class: This is a class module that manages facility information.
- Hazard class: This is a class module that manages disaster information such as fire.
- Agent class: This is a class module that manages agent information.

Among the five classes, the agent class is applied only to the evacuation simulation. In addition, the agent class is not part of BIM data but a separate algorithm built into the evacuation simulation body. Since this study establishes BIM data for efficient use of fire and evacuation simulations on construction sites, agent classes are excluded.

Space class

Space class has all the spatial information of the building that is the target of fire and evacuation simulation and is the class that should be considered first in the BIM during the design stage. The most basic BIM data for this class is information that can identify each floor. This information can be defined for each floor of the building, and it should include the number of floors, the difference in height between the floor and the ceiling of each floor, and the height value of the floor of each floor. Next, as a single space with a limited area, the zone, which is the upper level of information representing the structure of the building such as space and stairs, should be included in the BIM during the design. Also, most of the simulations use the grid concept in the body. Grid is a detailed space created at equal intervals on the x-axis and y-axis within the zone and boundary. In order to efficiently distinguish spatial grids in simulation, grid information must be included in BIM data. Grid information that can be included in BIM data includes grid number, x-coordinate, y-coordinate, grid vertical coordinate, and zone number to which the grid belongs. In order to use the BIM spatial information divided into these detailed classes in the fire and evacuation simulation algorithm, a link concept is required. A link provides information to connect between two zones, or to connect between zones and the different floors. For example, it becomes a mediator through which a fire or agent moves from one zone to another, such as a door, wall, or stair. The relationship of various variables of space class of BIM is conceptually summarized in the figure 4-4. Also, table 4-5 shows the variables of spatial information required in the fire evacuation simulation.

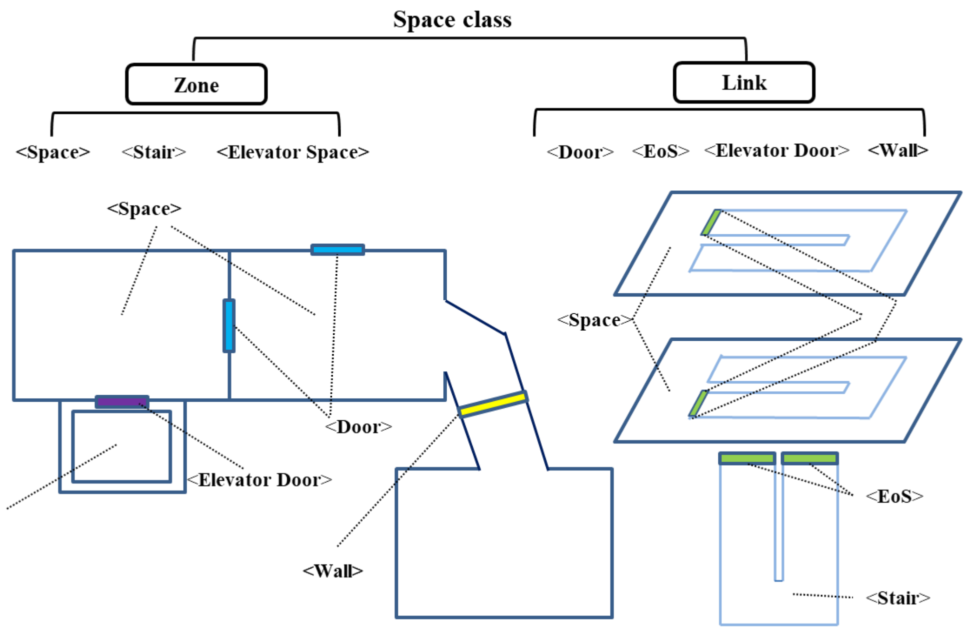

Figure 4-4 Relational diagram of space class

Table 4-5 List of space class variables for evacuation simulations

Category	Variable	Type/Class	Explanation
Floor	NO	Int	floor number (0 plus 1 from the lowest floor)
	Height	Float	The difference in height between the floor and the ceiling of the plan
	Elevation	Float	floor elevation of the plan
Zone	Type	Zone type	Zone type (Space/Stair), 0 for Space and 1 for Stair (example).
	ID	Int	Identification ID in BIM (SPACE_ID for Space, STAIR_NAME for Stair)
	NO	Int	Zone number (starting from 0)
	nGridX	Int	Number of grids in the x-axis direction
	nGridY	Int	Number of grids in the y-axis direction
	nGrid	Int	Number of Grids in Zone

	nLink	Int	Number of Links in Zone
	Height	Float	The vertical length between the floor and the ceiling of the zone
	Elevation	Float	The height value of the lowest floor of the zone
	Boundary	Polygon	Polygons that form the outer area
Grid	NO	Int	Number of grids
	NO_X	Int	X-direction number of grids
	NO_Y	Int	Y-direction number of grids
	ZoneNO	Int	Zone number to which the grid belongs
	Point	Int	Grid position coordinates (x,y)
	Z	Int	Grid vertical coordinates
Link	NO	Int	Number of Links
	FloorNO	Int	The number of the floor to which the link belongs
	nZone	Int	The number of zones that the link is in contact with
	CenPoint	Point	Center points of the link
	Width	Float	Link width
	Type	Link type	Type of Link (DoorLink, WallLink, StairLink, ElvLink)
	Elevation	Float	Link's elevation (same as Floor's elevation)
	V_ZoneNO	Int	Array of Zone numbers adjacent to the Link

The space information of the building required for the construction site fire and evacuation simulation analysis should be formed into the structure required by the simulation. In this study, the input data structure required for the fire evacuation simulation was set as the structural frame of the building shape information, and the necessary contents were arranged so that the IFC could read it. Table 4-6 shows the conversion relationship between the building data provided by IFC and the information required in the fire evacuation simulation. Current IFC data is used to define space information, and for accurate fire and evacuation analysis, it is necessary to clarify zone information. A zone must contain more than one space, and up to one floor can be considered as a zone. Also, zones must be connected by a separate structure such as a door. The operation of elevators and HVAC systems in buildings is not yet supported by BIM software, so users must manually input this information for fire and evacuation simulations. For example, the representation of HVAC and elevators in BIM software is possible, but it is difficult to input the operating method and details of each system required for building performance analysis simulation. In addition, there are several main principles for using IFC data as space information for fire and evacuation simulations. First, the terms, contents, and principles necessary for generating space information of a building should be written based on the simulation input file. Second, the coordinate values for expressing the shape of the structure should be unified based on the lower left point. Third, the floor must designate the final evacuation floor as a separate floor, and the rest as the floor of each floor. Fourth, since the door has a great influence on the evacuation simulation result, it is necessary to perform the simulation by checking whether the door is open or not.

Table 4-6 IFC to fire and evacuation simulation information transformation relationship

	IFC Information		Fire Simulation Require Information	
	Lowest level object	Attribute	Component	Item
ZONE	IfcSpace	Name	Zone	Name
STRUCTURE	IfcBuildingElement	Name	Types	Wall
				Floor
				Ceiling
				Roof
	IfcRelSpaceBoundary	PhysicalOrVirtualBoundary	Adjacent space	Surface
				Outdoor
	IfcCartesianPoint	Coordinates		Ground
	IfcRelSpaceBoundary	InternalOrExternalBoundary	Adjacent space environment	SunExposed
				WindExposed
				NoExposed
	IfcBuildingElement	Name	View Factor	0
				0.5
				1
	IfcCartesianPoint	Coordinates	Coordinate	X
				Y
				Z
			Properties	Roughness

	IfcPropertySingleValue	NominalValue		Thickness
				Conductivity
				Thermal Absorptance
DOOR	IfcBuildingElement	Name	Types	Window
				Door
				Glass Door
	IfcCartesianPoint	Coordinates	Surface information	Name
				Outside Boundary Condition Object
	IfcBuildingElement	Name	View Factor	0
				0.5
				1
	IfcCartesianPoint	Coordinates	Coordinate	X
				Y
				Z

In addition, when applying BIM to space input data in fire and evacuation simulations, it can help in simulation calculations. In the evacuation simulation, each occupant must have information about the space in the building to reach their destination, and it is the distance potential field that provides information about the space (Anvari et al. 2014; Kretz et al. 2010). The distance potential field calculates the distance to a specific destination and reflects the arrangement of walls, doors, stairs, columns, etc. that physically form an indoor space as a geometric element (Kretz 2010). This is widely used in many evacuation simulations such as Simulex and Egress developed previously (Rogsch and Klingsch 2012). However, this method has limitations. First, if the space is large or complex, the number of cells to be calculated increases rapidly, which increases the computation time. In addition, there is a limit to the immediate recalculation if the destination needs to be updated according to the evacuation simulation and fire simulation progress. The use of BIM data proposed in this study can improve these limitations. In the case of BIM data, each space can have its own number and data. With a clear separation of each space, it is possible to calculate by dividing a number of small spaces in the evacuation and fire simulations and then recombining them. This is called a Hierarchical

Distance Potential Field. To use this methodology, a space class including space geometry must be clearly defined, and this can be implemented using BIM data. In addition, the Hierarchical Distance Potential Field can provide more visible results after the simulation is complete. This can provide users with more meaningful information compared to traditional result files that simply derive fire and evacuation completion times. For example, the user can check the density of each space through the Hierarchical Distance Potential Field, which can be used to consider installing additional doors or evacuation equipment.

In BIM, information on objects included in space is more detailed than in 2D plans. BIM data containing detailed spatial information can be used to improve the accuracy of simulations. For example, it is difficult to apply the exact height of objects such as obstacles in simulations using 2D plans. However, since the BIM data includes the height values and materials of objects in the space, more accurate simulations are possible by using them. In addition to avoiding obstacles in traditional simulations, various movements such as crossing obstacles or windows can be added to the simulation. This can increase the resemblance to real fire and evacuation situations in buildings, which improves the accuracy of the fire and evacuation simulation results and gives the simulation reliability. In addition, after completing the fire and evacuation simulation using BIM data, it provides information to improve the fire safety of the building through the relocation of objects in the identified hazardous area.

Path class

The path class is a module that creates a topology based on the basic spatial shape of a building and processes the path and distance information of the entire building. Depending on the factors affecting the path of fire and evacuation, it can be created multiple times. It is created at the beginning of the simulation, but if the factors affecting the path change during the simulation, it can be recreated and changed each time. Since this class is created based on the space class, many variables correspond to the space class. Through table 4-7, the author shows the BIM data variables required to implement the path class to use evacuation simulation. Variables of path class are not converted directly from BIM data to path class but are created based on space class.

Table 4-7 List of path class variables to be included in the BIM proposed in this study

Category	Variable	Type/Class	Explanation
Path	Dest	Int	Final destination location
	EXIT	Int	Exit in each space unit
	Linked	Boolean	Is the Zone connected to another Zone through a Link?
	NearExit	Boolean	Is this Zone adjacent to the Link corresponding to Dest of this Path?
	LinkDist	Float	Shortest distance value to Dest after calculating the shortest path
	NestLinkNO	Int	When calculating the shortest path, the number of the next link that should be proceeded from this link

In the fire and evacuation simulation using a 3D plan based on BIM data, it is possible to provide a result file about the occupant's movement path. It is effective in predicting and responding to crowd paths and bottlenecks through more intuitive visualization when using a 2D plan. When reviewing a path in a BIM-based drawing, the user can figure out in advance the space where the path is not activated due to incorrect plans. In addition, the analysis of the evacuation path of the occupants can be reference information for determining the entry route of firefighters. In this study, the variables of the path class were verified using the BIM-based evacuation simulation that is being developed as a prototype by Korea's National Fire Agency. This simulation creates topology data based on the basic spatial shape of a building extracted from BIM and performs evacuation routes and simulations. A path is created considering the variables proposed in Table 4-7. Paths are created at the beginning of the simulation, but if the factors affecting the path (fire spread and congestion) change during the simulation, the path is recalculated each time. Figure 4-5 is an example of providing path information in the BIM based simulation developed by Korea's National Fire Agency.

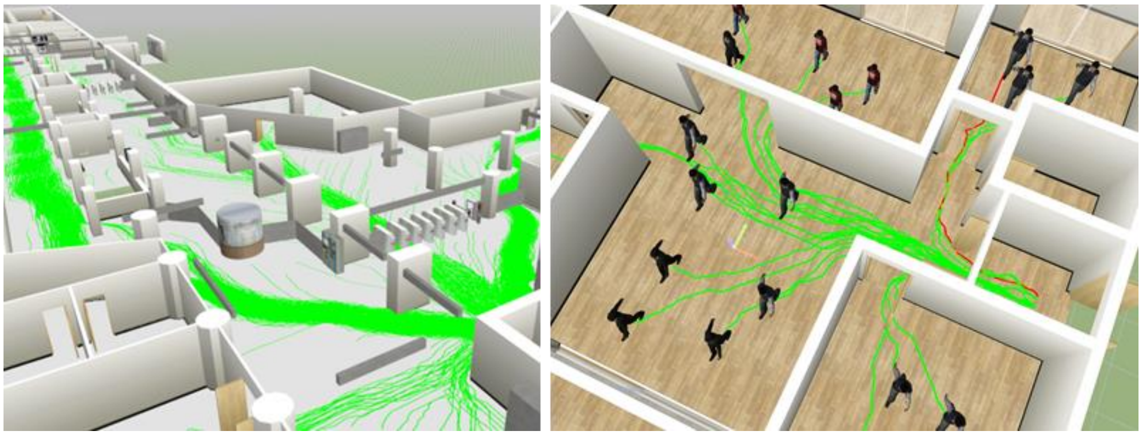

Figure 4-5 Example of path information of BIM-based evacuation simulation proposed in this study

In addition, since BIM contains information on the number of floors in each space, it is more suitable for setting various paths in the evacuation simulation of a high-rise building. Since the existing evacuation simulations are calculated separately for each floor, there is a limit to the accuracy of the evacuation simulation for high-rise buildings. BIM data includes information on the number of floors and stairs in each space, so it is possible to calculate multiple evacuation paths for high-rise buildings, which can be used to increase evacuation simulation accuracy.

Facility class

Facility Class is a class that manages facilities related to fire and evacuation. Current BIM software, including RevitMEP, lack design support for facility objects such as elevators and HVAC systems in buildings. The visual expression of the elevator and HVAC system is possible, but it is impossible to input specific data of each system's operation information required for fire and evacuation simulation evaluation. Therefore, when converting IFC data to fire and evacuation simulation input files, facility data must be added. Also, in order to perform the simulation, some simulation parameters must be manually entered according to each analysis program. In most current simulators, the facility class is implemented only for elevators. Therefore, in this study, variables that can be included in BIM data were extracted by limiting the facility class to the elevator. Through additional research, if various facility algorithms such as evacuation guidance lights, evacuation alarms, and fire detectors are added to the simulation, the structure and data of the facility class can be added. The facilities have independent operation algorithms and are implemented in such a way that they operate when certain conditions are met during simulation. Table 4-8 shows the BIM data variables required by the facility class.

Table 4-8 List of elevator class variables to be included in the BIM proposed in this study

Category	Variable	Type/Class	Explanation
Facility	ID	Int	Elevator's ID
	BottomPlan	Int	The lowest floor of the elevator
	TopPlan	Int	The top floor of the elevator
	nAgent	Int	Total number of agents using this elevator
	V_AgentNO_OnBoard	Int	Array of agent numbers on board the elevator
	ElvDoorID	Int	Corresponding ElvDoor ID
	LinkNO	Int	Link number corresponding to ElvDoorID
	Timer_Elv	Int	Timer for elevator door control
	Speed	Float	Elevator vertical movement speed (cm/sec)
	Z	Float	Elevator floor elevation vertical coordinates
	Inner_Width	Float	Left and right width of elevator

Inner_Depth	Float	The vertical width of the elevator (inward from the entrance)
CenPoint	Point	Elevator center coordinates

Hazard class

A hazard class is a class for generalizing disaster data such as fire and smoke. Although only fire and smoke data are dealt with in this study, various disaster data may be added later. Similar to the path class, in order to effectively calculate the disaster data of the hazard class, BIM data including the space class is required. In addition, if the hazard data is designed to be used in the form of a grid, external disaster simulation results can be applied. Table 4-9 shows the variables required to efficiently use hazard data in the relevant simulations.

Table 4-9 List of hazard class variables to be included in the BIM proposed in this study

Category	Variable	Type/Class	Explanation
Hazard	nSec	Int	Duration of hazard data
	nHazardGrid	Int	Number of hazard grids
	V_HazardGrid	Vector	Hazard grid array
	V_Sec	Vector	Time (seconds) array
	V_Temper	Vector	Temperature
	V_CO	Vector	Array of carbon monoxide numbers
	V_CO2	Vector	Array of carbon dioxide figures
	V_SOOT_Visi	Vector	Visibility

By utilizing BIM-based plans, it is possible to provide results by linking fire simulation and evacuation simulation. Through BIM-based simulation, users can more intuitively visualize the fire situation and progress. In addition, a function to determine the success of evacuation due to fire based on the current location of the occupant can be added to the simulation. Calculating the temperature and smoke level of a certain space through fire simulation can affect the moving agent. It can be confirmed in the simulation how the progress of the fire affects the evacuation of the occupants. This has the ability to expand to incorporate the criteria for determining the tenability limit in a building fire in future studies. The criteria for determining the tenability limit of a building fire include heat exposure and suffocation, and the safety of occupants can be confirmed by reflecting the relevant criteria in the input data in the simulation.

In this study, the fire variables were applied to the evacuation simulation using the BIM-based evacuation simulation being developed as a prototype by Korea's National Fire Agency. This evacuation simulation links the fire information from the external fire simulation, applies it to the evacuation algorithm, and visualizes the results to provide the user. In the current prototype stage, only FDS was linked during fire simulation, but if the input data format is matched, various fire analysis programs are applied. In addition, the BIM-based evacuation simulation has height information for each space, unlike the 2D plan-based evacuation simulation. Based on this height information, it is possible to measure the fire risk for each space in the evacuation simulation. In this simulation, the algorithm is designed so that when the smoke comes down to a height of 1.8m, which is the limited smoke layer standard of the Korean fire safety regulation, the space is set as an area where evacuation is impossible. By visualizing the data provided by the fire simulation, such as the temperature, smoke amount, and CO concentration of each space, users can intuitively analyze the fire risk of a space. Figure 4-6 is an example of an evacuation simulation design considering a limited smoke layer (1.8m) and visualization of fire information in the simulation.

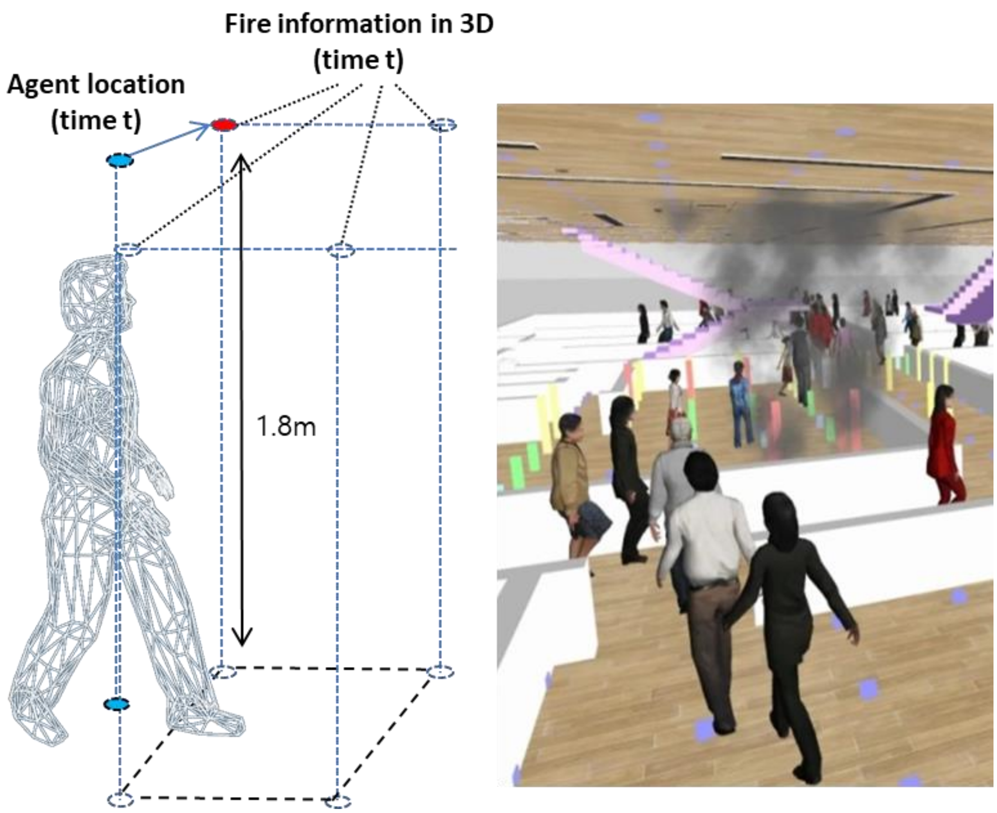

Figure 4-6 Example of evacuation simulation design considering limit smoke layer

BIM validation module for fire and evacuation simulation

The BIM validation module is a module that reviews whether BIM data created from BIM software is extracted in a form suitable for fire and evacuation analysis or not. This is an important part to improve the reliability of BIM-based fire and evacuation simulation analysis results. Therefore, in this study, the items to be verified in the BIM validation module were proposed, and examples of detailed items for each division were presented. The verification category of the BIM validation module can be divided into the existence or non-existence of objects, attribute information, and object suitability. Table 4-10 shows the classification of BIM validation module verification categories.

Table 4-10 Classification of BIM validation module validation categories

Verification category	Contents
Object existence	Determine the existence of basic objects necessary for fire and evacuation analysis
Attribute information	Review the validity of attribution information required for fire and evacuation analysis
Object suitability	Evaluate the suitability of modeling state of building objects from the perspective of fire and evacuation analysis

In addition, rules can be defined to implement IFC data verification for fire and evacuation performance evaluation. After classifying items for evaluating fire and evacuation performance, the rule can assign verification contents and entities. Table 4-11 shows the rules for the existence of doors connecting each space, which are examples of rules for fire and evacuation performance evaluation. Rules can be added in simulation types and stages, and BIM data can be verified with a rule checker using standardized rules.

Table 4-11 Examples of IFC data validation rules proposed in this study

Division	Item	Details	BIM guidelines
IFC	Object Existence	Door	-
Verification Content	Check whether the door object of the target building exists		
Rule Description	Outputs True if the IfcRelSpaceBoundary containing the IfcDoor of the target building exists, and False if it does not exist.		
Entity	IFCRelSpaceBoundary	IfcDoor	-

BIM-based fire and evacuation simulation user process

The user process of the BIM-based fire and evacuation simulation framework proposed in this study is classified into five steps.

- Step 1: BIM modeling

- The target building for fire and evacuation performance evaluation is modeled through BIM software and extracted in IFC, an international standard format. When modeling, the BIM modeling guideline established in this study or related simulations should be followed.

- Step 2: BIM model validation

- It is necessary to review whether the IFC file generated by the BIM software is modeled according to the relevant guidelines. If it is reviewed as an inappropriate file in the BIM validation module, BIM modeling must be performed again based on the related rules.

- Step 3: Create a project

- Create a project within the fire and evacuation simulation and input the converted input file from the relevant converter.

- Step 4: Fire and evacuation analysis modeling

- In addition to the information defined in the BIM model, attribute information, including space, path, facility, hazard, and agent information, is required for building fire and evacuation analysis.

- Step 5: Perform simulation and visualize results

- When all attribute information required for fire and evacuation analysis is entered, the simulation engine is started. To perform the simulation, it can select the simulation period and the type of analysis model, and the simulation analysis time varies depending on the size of the target building. The simulation results can be visualized as a 3D spatial model as well as text output such as an excel file, and the simulation process can be provided as a video.

Importance of fire and evacuation simulation through BIM

Most of the existing fire and evacuation simulations are performed based on two-dimensional plans. For more accurate results of the fire and evacuation simulator, it is necessary to share and use the same spatial information based on 3D spatial information at the same time. When using two-dimensional spatial data in most of the relevant simulations, additional drawing data cleaning is required. In addition, if the clear BIM data required to use the simulation is not provided, the user has to redraw the 3D spatial information. In the fire and evacuation simulation, it is important to accurately input the information of the real space, and the most efficient way to check this accuracy is for the user to visually compare the real space with the spatial information implemented in the simulation. This can be visually compared in buildings designed with BIM data and is a starting point to increase the accuracy of simulations. The current architectural design is actively utilizing the advantages of BIM, which can provide information, such as 2D/3D models, required materials, construction, and construction period at once. BIM can be used as basic input data for fire and evacuation simulations considering the usability, work efficiency, and fire safety application range. Therefore, it is essential to secure interoperability with BIM technology in advance for fire safety on construction sites.

In addition, the use of BIM data can be extended to various fields of fire safety as well as fire and evacuation simulations. Firefighter training simulation using BIM-applied building data can be proposed. Traditional firefighter training is limited to using documented materials such as 2D drawings or using related game programs. (Lebram et al. 2009). This training program has a limitation that the training is conducted in a virtual building rather than a building within the area where firefighters can be dispatched. In addition, firefighter training using 2D plans has a disadvantage that it is difficult for firefighters to properly understand the overall shape and structure of the building. However, the effectiveness of the training can be improved when the training is conducted on real buildings in each fire department's area designed with BIM. Since BIM data provides a 3D space for a building, it can be applied to firefighters' fire suppression training using virtual reality. As the types of buildings designed with BIM are expected to increase in the future, it will be necessary to develop a BIM-based firefighter training program (Huang et al. 2021). Figure 4-7 is a BIM-based firefighter training program developed by Gyeongbuk fire service headquarters. This simulation includes the entire process from the placement of a fire engine on the outside of a building to fire suppression indoors. Since this simulation uses the already established BIM data of each fire department's control area, there is no need for additional drawing work for the simulation. By using BIM to provide firefighters with a 3D building modeling screen that is similar to a real building, the understanding of firefighters' training is improved.

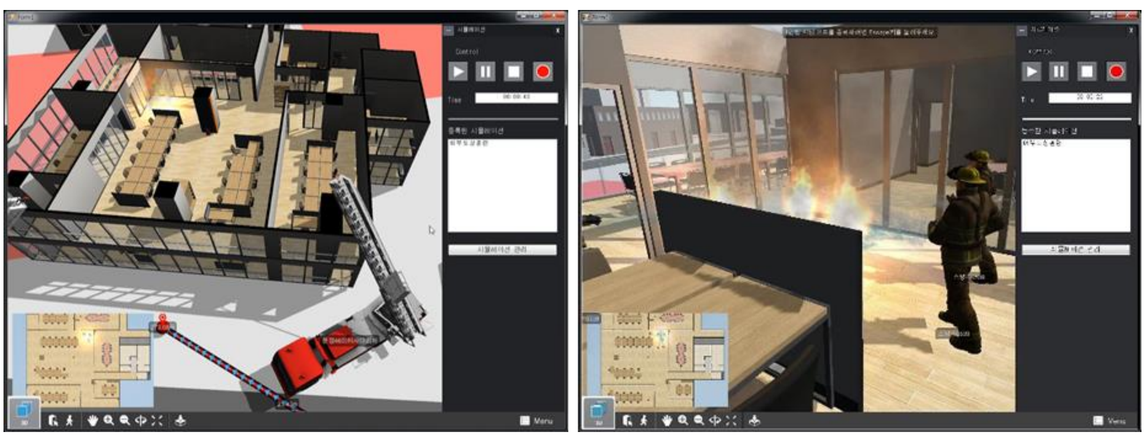

Figure 4-7 BIM-based firefighter training program developed by Gyeongbuk fire service

CONCLUSION AND RECOMMENDATIONS

Researchers have rarely focused on fire and evacuation safety on construction sites due to its low occurrence compared with other types of accidents. Previous studies that analyzed fire safety on construction sites have mainly focused on traditional fire hazard materials and evacuation simulations. However, these studies have not yet explored specific methods for improving fire safety or analyzing fire and evacuation simulations for construction sites. In addition, BIM has been rarely considered as means to analyze fire safety on construction sites . This study provides information for the efficient use of BIM data that can improve fire safety on construction sites. By establishing BIM data that should be considered when using fire safety simulations on construction sites, it provides information necessary to be created in BIM during design stage. This study contributes to the body of knowledge by providing specific information for applying BIM to fire and evacuation simulations.

To analyze fire safety on construction sites, new approaches such as the application of BIM data were introduced and used in this study. In order to prevent fire accidents on construction sites, it is necessary to considering fire and evacuation safety during BIM content development during the design. This study provides data overview and details for BIM models considering construction site fire safety analysis. In this study, BIM data required for fire safety analysis of construction sites were classified into five main classes, and detailed variables and data types to be included in each class were explored. In addition, in order to efficiently apply BIM data to fire and evacuation simulations, a BIM validation module was provided. Through the BIM validation module suggested in this study, the suitability of BIM data can be checked before running the simulation. Users can add rules to check BIM data used for fire and evacuation simulations. The user process of the BIM-based simulation framework provided in this study is divided into five steps. The five steps consist of BIM modeling, BIM model validation, creating a project, fire and evacuation analysis modeling, performing simulation and visualizing results. In addition, it was emphasized that the use of BIM data in this study can be extended to various fields of fire safety as well as fire and evacuation simulations. Unlike 2D plans, BIM has the advantage of visualizing a building three-dimensionally, and by developing various safety education programs using this, it is possible to improve the understanding of trainees. By establishing relevant BIM data for effective analysis of fire and evacuation safety on construction sites, more effective fire safety measures such as manuals and regulations can be developed.

The methodology and results chosen for this study presented several limitations to the analysis. The data collected in this study has a limited target, and various data can be added through future research. The current study analyzed the structure and details of BIM data for fire safety analysis of construction sites were presented, but in future studies, variables included in the main class of BIM data can be added. Therefore, future research should suggest ways to improve fire safety by introducing and analyzing BIM-based fire and evacuation simulations on construction sites and buildings and verifying the results. This research focuses on the use of BIM-based simulation to improve fire safety on construction sites. However, due to the characteristics of the construction site, many fire safety facilities may not be completely installed depending on the progress of the construction. This causes a difference between the BIM data and the design of the actual construction site, which may affect the simulation results. To compensate for this, it is necessary to include accurate schedule information in BIM data. Although 4D information including schedules can be additionally applied to BIM data, there is a lack of research on extracting BIM data for each stage of construction using this. In addition, since the schedule of each construction stage can change depending on the situation on the construction site, it is necessary to prepare a plan to continuously update these changes through future research. If BIM data is extracted and used according to each stage of construction, the reliability of simulation results can be improved. BIM basically expresses the shape information of a building. There is a limit to the inability to include numerous materials or dangerous activities on the construction site in BIM data. In particular, the storage of flammable materials is an important factor in the risk of fire on construction sites, but these storage materials cannot be included in BIM because they are not the shape information of the building. In addition, dangerous construction activities that can cause fires on construction sites cannot be applied to BIM data because they are not data in terms of building design. Through further research, an expanded framework is needed to define various activities on construction sites that can cause fires and include them in fire and evacuation simulations. Fire risk factors that are not included in the building shape information should be separately managed and reflected in fire safety evaluation and simulation.

CHAPTER V

CONCLUSION AND RECOMMENDATIONS

SUMMARY OF RESULTS AND CONTRIBUTIONS

The main objectives of this dissertation are to (a) find the fire accident factors on construction sites using web crawling and deep learning approach, (b) use statistical methods such as path analysis to identify factors influencing fire safety on building construction sites and (c) provide BIM data class classification and guidelines for efficiently applying BIM to fire and evacuation simulations on construction sites.

The research presented in Chapter 2 contributed to the body of knowledge in various ways. Based on the previous research, it was hard to find related research on the relationship between the factors of fire accidents on construction sites. First, the author identified the reason why studies of fire safety for construction sites are insufficient. Researchers have rarely focused on fire research on construction sites due to its low occurrence compared with other accident types. Researchers also have not adequately analyzed the main factors of fire accidents on construction sites and the effects of fire accidents. Second, various data mining techniques were used to define the factors of fire accidents on the construction site and to understand the relationship between the factors. By applying web crawling and deep learning techniques, keywords of fire safety on construction sites were analyzed to identify the main causes of fire accidents. Third, this study provided that building and fire code violations, lack of regular inspections on construction sites, and an incomplete fire safety system were major factors in fire accidents on construction sites. This is not a result of the careless and inappropriate behavior of labor on construction sites, but the main factors that can be prevented with regulations and on-site fire safety protection systems. By finding the main reasons for fire accidents on construction site and exploring the relationship between the fire factors, this study provides meaningful contributions to developing fire safety regulations that consider the characteristics of construction sites.

The analysis presented in Chapter 3 presents a method to identify fire factors in construction sites based on the perceptions of firefighters. This study provides information on the relationship between the importance of fire safety and the factors influencing fire in the construction stage which has not been studied before. Most building safety codes and regulations are established for buildings under the occupancy stage and the construction stage is inadequately considered. The results suggest that fire risk on the construction stage is as important as a completed building, and in particular, firefighters answered that an effective safety system on the construction site is important. There were noticeable differences in the perceived importance of factors affecting fire safety on construction sites depending on the fire officer's job specialty. Through the structural equation modeling, this chapter found that building factors had the greatest direct effect on fire factors concerning fire safety on construction sites. Although human factors did not have a strong direct relationship with fire factors, the human factors had an indirect impact on fire-related factors through the human factor as a mediator to achieve fire safety.

In Chapter 4, the author provided a method to improve the accuracy of fire and evacuation simulations on construction sites through the definition of related BIM data. The goal of this study was to effectively apply BIM to fire and evacuation simulations of construction sites. To analyze fire safety on construction sites, new approaches such as the application of BIM were used in this chapter. This suggests that the application and education of a new methodology to improve fire safety on construction sites is necessary. To apply BIM to fire and evacuation simulations of construction sites, the author classified the data into five main classes. It was constructed based on the information required in the calculation stage of the current fire and evacuation simulation, and also considered future extendability. Therefore, this study recommends improving the fire and evacuation safety of the construction site by using BIM for the fire evacuation simulation of construction sites.

LIMITATIONS AND FUTURE RESEARCH

This dissertation was mainly composed of three purposes, and different research targets and methodologies are applied to each chapter. The research targets include web data on fire accidents on construction sites, investigation of firefighters' perceptions of fire accident factors, and classification of BIM data for effective application to fire and evacuation simulations. Web crawling, deep learning, word embedding, survey, path analysis, and BIM data classification were used as the methods for this study. The research in each chapter had limitations that need to be addressed in future studies.

In Chapter 2, the factors related to fire safety on the construction site were defined and the relationship between the factors was identified. The methodology for this study has some limitations. Since the data collected in this study is limited to media companies, it is necessary to collect data from various targets through future research. Although this study presented the relationship and similarity between major keywords of fire accidents on construction sites, it is necessary to explore the relationship between sub-keywords belonging to each keyword. Applying more advanced data collection and analysis techniques to improve reliability can reinforce the results of this study.

Chapter 3 presented some limitations to the methodology and results chosen for this study. The current study focuses on the context of construction sites located in densely populated areas. The value of this study will be enhanced if further studies on non-populated areas are added through future research. This study also only focuses on the perception of firefighters. As different occupational groups may have different perceptions of fire on construction sites, future research should consider other relevant stakeholders. A good example would be construction workers or designers on a construction site. In addition, it is necessary to investigate what kind of building type the firefighters responded to when they participated in this survey. Depending on the type of construction site, there may be differences in the perception of survey participants.

Therefore, future studies will be able to deal with construction site fire safety on a larger scale by considering the above-mentioned limitations.

In Chapter 4, the research on the application of BIM data for fire safety on construction sites was lacking, so the study was limited to the application of BIM data for fire and evacuation simulations on construction sites. Future research should be extended to the application of BIM data to improve fire safety from various perspectives, including fire and evacuation simulations. The current study analyzed the class and details of BIM for fire and evacuation safety analysis of construction sites were presented. The results of this study make it evident that it is very important to effectively apply BIM to fire and evacuation simulations, and more future research needs to be conducted in this area. In future research, it is possible to add the class and variables included in the class required for the simulation proposed in this study. In addition, due to the characteristics of the construction project, many safety facilities may not be fully installed depending on the stage of the construction project. This causes a difference between the actual construction site and the BIM data, which may affect the fire and evacuation simulation results. To fill the knowledge gap, it is necessary to add construction schedule information in BIM data. If BIM data is extracted according to each construction stage through future research and applied to the simulation, the reliability of fire and evacuation simulation results on the construction site will be improved. Through future research, an expanded BIM-based framework that can define various factors and construction activities on sites that can cause fires and include them in fire and evacuation simulations are needed.

REFERENCES

Abbas, M., Mneymneh, B. E., and Khoury, H. (2018). "Assessing on-site construction personnel hazard perception in a Middle Eastern developing country: An interactive graphical approach." *Safety Science*, 103, 183-196.

Abdelhamid, T. S., and Everett, J. G. (2000). "Identifying root causes of construction accidents." *Journal of construction engineering and management*, 126(1), 52-60.

Abdullah, D., and Wern, G. C. M. "An analysis of accidents statistics in Malaysian construction sector." *Proc., International Conference on E-business, Management and Economics*, IACSIT Press Honk Kong, 1-4.

Ahn, B.-J., Jung, K.-H., Lee, J.-S., and Rhim, J.-K. (2018). "A Study on the Risk of Organic Solvents for Underground Area under Construction Site through a Fire Accident Case." *Journal of the Korean Institute of Gas*, 22(1), 9-17.

Ahn, C.-S., Bang, B.-H., Kim, M.-W., James, S. C., Yarin, A. L., and Yoon, S. S. (2019). "Theoretical, numerical, and experimental investigation of smoke dynamics in high-rise buildings." *International Journal of Heat and Mass Transfer*, 135, 604-613.

Aires, M. D. M., Gámez, M. C. R., and Gibb, A. (2010). "Prevention through design: The effect of European Directives on construction workplace accidents." *Safety science*, 48(2), 248-258.

Ali, A., Kamaruzzaman, S., and Sing, G. (2010). "A Study on causes of accident and prevention in Malaysian construction industry." *Journal of Design+ Built*, 3(1), 95-104.

Ali, K. N., Alhajlah, H. H., and Kassem, M. A. (2022). "Collaboration and Risk in Building Information Modelling (BIM): A Systematic Literature Review." *Buildings*, 12(5), 571.

Alkaissy, M., Arashpour, M., Zeynalian, M., and Li, H. (2022). "Worksite Accident Impacts on Construction and Infrastructure: Nondeterministic Analysis of Subsectors and Organization Sizes." *Journal of Construction Engineering and Management*, 148(6), 04022023.

Anderson, J. C., and Gerbing, D. W. (1988). "Structural equation modeling in practice: A review and recommended two-step approach." *Psychological bulletin*, 103(3), 411.

Anvari, B., Daamen, W., Knoop, V. L., Hoogendoorn, S. P., and Bell, M. G. (2014). "Shared space modeling based on social forces and distance potential field." *Pedestrian and evacuation dynamics 2012*, Springer, 907-916.

Ayhan, B. U., and Tokdemir, O. B. (2019). "Predicting the outcome of construction incidents." *Safety science*, 113, 91-104.

Bag, S., Kumar, S. K., and Tiwari, M. K. (2019). "An efficient recommendation generation using relevant Jaccard similarity." *Information Sciences*, 483, 53-64.

Baozhi, X. G. W. L. C. (2001). "Review and Perspective in the Study of Human Behavior during Evacuation from Building Fire [J]." *China Safety Science Journal*, 3.

Bentler, P. M., and Chou, C.-P. (1987). "Practical issues in structural modeling." *Sociological methods & research*, 16(1), 78-117.

BLS (2019). "Establishment Specific Injury and Illness Data." <https://www.osha.gov/Establishment-Specific-Injury-and-Illness-Data>.

Campbell, D. T., and Cook, T. D. (1979). *Quasi-experimentation: Design & analysis issues for field settings*, Rand McNally College Publishing Company Chicago.

Campbell, R. B. (2017). *Fires in structures under construction, undergoing major renovation, or being demolished*, National Fire Protection Association Quincy, MA.

Carbonari, G., Stravoravdis, S., and Gausden, C. (2018). "Improving FM task efficiency through BIM: a proposal for BIM implementation." *Journal of Corporate Real Estate*.

Carvalho, J. P., Almeida, M., Bragança, L., and Mateus, R. (2021). "Bim-based energy analysis and sustainability assessment—Application to portuguese buildings." *Buildings*, 11(6), 246.

Caton, S. E., Hakes, R. S., Gorham, D. J., Zhou, A., and Gollner, M. J. (2017). "Review of pathways for building fire spread in the wildland urban interface part I: exposure conditions." *Fire technology*, 53(2), 429-473.

Chan, M. (2011). "Fatigue: the most critical accident risk in oil and gas construction." *Construction Management and Economics*, 29(4), 341-353.

Chen, H., Hou, L., Zhang, G. K., and Moon, S. (2021). "Development of BIM, IoT and AR/VR technologies for fire safety and upskilling." *Automation in Construction*, 125, 103631.

Chen, J., Liu, C., Meng, Y., and Zhong, M. (2021). "Multi-Dimensional evacuation risk evaluation in standard subway station." *Safety science*, 142, 105392.

Cheng, Y., Liu, D., Chen, J., Namilae, S., Thropp, J., and Seong, Y. "Human behavior under emergency and its simulation modeling: a review." *Proc., International Conference on Applied Human Factors and Ergonomics*, Springer, 313-325.

Choi, J., and Lee, S.-W. (2020). "Improving FastText with inverse document frequency of subwords." *Pattern Recognition Letters*.

Choi, J., Lee, S., and Kim, I. (2020). "Development of Quality Control Requirements for Improving the Quality of Architectural Design Based on BIM." *Applied Sciences*, 10(20), 7074.

Church, K. W. (2017). "Word2Vec." *Natural Language Engineering*, 23(1), 155-162.

Cooper, M. D., and Phillips, R. A. (2004). "Exploratory analysis of the safety climate and safety behavior relationship." *Journal of safety research*, 35(5), 497-512.

Costello, A. B., and Osborne, J. (2005). "Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis." *Practical assessment, research, and evaluation*, 10(1), 7.

D'Haen, J., Van den Poel, D., Thorleuchter, D., and Benoit, D. F. (2016). "Integrating expert knowledge and multilingual web crawling data in a lead qualification system." *Decision Support Systems*, 82, 69-78.

Dagan, D., and Isaac, S. (2015). "Planning safe distances between workers on construction sites." *Automation in construction*, 50, 64-71.

De-Ching, H., Shen-Wen, C., Chien-Hung, L., Po-Ta, H., Yi-Ting, S., and Huei-Ru, S. (2011). "A study for the evacuation of hospital on fire during construction." *Procedia Engineering*, 11, 139-146.

DeCoster, J. (1998). "Overview of factor analysis."

Dimyadi, J., Solihin, W., and Amor, R. "Using IFC to support enclosure fire dynamics simulation." *Proc., Workshop of the European Group for Intelligent Computing in Engineering*, Springer, 339-360.

Doglione, R., and Piccinini, N. (2022). "Concurrent design error, poor welding, and code inadequacy cause jet-FIRE on a topping column." *Process Safety Progress*.

Farghaly, K., Soman, R. K., Collinge, W., Mosleh, M. H., Manu, P., and Cheung, C. M. (2022). "Construction safety ontology development and alignment with industry foundation classes (IFC)." *Journal of Information Technology in Construction*, 27, 94-108.

Feng, K., and Lin, N. (2022). "Modeling and analyzing the traffic flow during evacuation in Hurricane Irma (2017)." *Transportation research part D: transport and environment*, 110, 103412.

Folk, L. H., Kuligowski, E. D., Gwynne, S., and Gales, J. A. (2019). "A provisional conceptual model of human behavior in response to wildland-urban interface fires." *Fire technology*, 55(5), 1619-1647.

Gaur, S., and Tawalare, A. (2022). "Investigating the role of BIM in stakeholder management: Evidence from a metro-rail project." *Journal of Management in Engineering*, 38(1), 05021013.

Gerges, M., Demian, P., and Adamu, Z. (2021). "Customising evacuation instructions for high-rise residential occupants to expedite fire egress: Results from agent-based simulation." *Fire*, 4(2), 21.

Guo, B. H., Yiu, T. W., and González, V. A. (2016). "Predicting safety behavior in the construction industry: Development and test of an integrative model." *Safety science*, 84, 1-11.

Guy, I., Schwartz, I., and Radinsky, K. (2019). "Search system for providing web crawling query prioritization based on classification operation performance." Google Patents.

Haghani, M. (2020). "Optimising crowd evacuations: Mathematical, architectural and behavioural approaches." *Safety science*, 128, 104745.

Hallowell, M. R., and Gambatese, J. A. (2009). "Activity-based safety risk quantification for concrete formwork construction." *Journal of Construction Engineering and Management*, 135(10), 990-998.

Hallowell, M. R., and Gambatese, J. A. (2009). "Construction safety risk mitigation." *Journal of Construction Engineering and Management*, 135(12), 1316-1323.

Hamid, A. R. A., Yusof, W. Z. W., and Singh, B. S. B. J. (2003). *Hazards at construction sites*, Universiti Teknologi Malaysia.

Hardison, D., Behm, M., Hallowell, M. R., and Fonooni, H. (2014). "Identifying construction supervisor competencies for effective site safety." *Safety science*, 65, 45-53.

He, J. (2021). "Earthquake evacuation simulation of multi-story buildings during earthquakes." *Earthquake Spectra*, 37(1), 95-113.

Hindocha, E., Yazhiny, V., Arunkumar, A., and Boobalan, P. (2019). "Short-text Semantic Similarity using GloVe word embedding."

Hinze, J., and Raboud, P. (1988). "Safety on large building construction projects." *Journal of Construction Engineering and Management*, 114(2), 286-293.

Hong, S.-H., Lee, S.-K., and Yu, J.-H. (2019). "Automated management of green building material information using web crawling and ontology." *Automation in Construction*, 102, 230-244.

Hosseinian, S. S., and Torghabeh, Z. J. (2012). "Major theories of construction accident causation models: A literature review." *International Journal of Advances in Engineering & Technology*, 4(2), 53.

Huang, M., Ninić, J., and Zhang, Q. (2021). "BIM, machine learning and computer vision techniques in underground construction: Current status and future perspectives." *Tunnelling and Underground Space Technology*, 108, 103677.

Huang, X., Li, H., Li, X., and Zhang, L. "Fire numerical simulation analysis for large-scale public building in 3D GIS." *Proc., IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 7522-7525.

Hui, L., Yongqing, W., Shimei, S., and Baotie, S. (2012). "Study on safety assessment of fire hazard for the construction site." *Procedia Engineering*, 43, 369-373.

Ibrahim, A. M., Venkat, I., and De Wilde, P. (2019). "The impact of potential crowd behaviours on emergency evacuation: an evolutionary game theoretic approach." *Journal of Artificial Societies and Social Simulation*, 22(1).

Ingason, H., Lönnermark, A., Frantzich, H., and Kumm, M. (2010). "Fire incidents during construction work of tunnels."

Jafari, M., Gharari, M., Ghafari, M., Omidi, L., Kalantari, S., and Asadolah-Fardi, G. (2014). "The influence of safety training on safety climate factors in a construction site." *International Journal of Occupational Hygiene*, 6(2), 81-87.

Jeon, G., and Hong, W. (2007). "Characteristic Changes in an Evacuation of Ecacuees by Using Phosphorescent Guidance Equipment; though the Simulated Escape Experiment." WCOGI.

Jeong, M.-J., Lee, M.-G., and Ham, E.-G. (2014). "Assessment of Fire Evacuation Safety for Building Construction." *Journal of the Korean Society of Safety*, 29(6), 119-124.

Kallada Janardhan, R., and Hostikka, S. (2019). "Predictive computational fluid dynamics simulation of fire spread on wood cribs." *Fire Technology*, 55(6), 2245-2268.

Kannan, S., Gurusamy, V., Vijayarani, S., Ilamathi, J., Nithya, M., Kannan, S., and Gurusamy, V. (2014). "Preprocessing techniques for text mining." *International Journal of Computer Science & Communication Networks*, 5(1), 7-16.

Kehui, X. (2011). "On the Construction Site of the Fire Hazards and Management Measures [J]." *China Public Security (Academy Edition)*, 2.

Khatua, A., Khatua, A., and Cambria, E. (2019). "A tale of two epidemics: Contextual Word2Vec for classifying twitter streams during outbreaks." *Information Processing & Management*, 56(1), 247-257.

Kim, H.-s., Jang, S.-c., and Joo, J.-g. (2019). "Preventive Priority Methods Based on the Analysis of Fire Accident Causes in Construction Site." *Journal of the Korea Institute of Construction Safety*, 2(2), 50-55.

Kim, H., Lee, J.-K., Shin, J., and Choi, J. (2019). "Visual language approach to representing KBimCode-based Korea building code sentences for automated rule checking." *Journal of Computational Design and Engineering*, 6(2), 143-148.

Kim, J.-S., and Kim, B.-S. (2018). "Analysis of fire-accident factors using big-data analysis method for construction areas." *KSCE Journal of Civil Engineering*, 22(5), 1535-1543.

Kim, J., Shan, Y., Kim, S., Song, D., Park, H., and Bang, C. (2021). "Factors Influencing Fire Safety on Building Construction Sites: A Fire Officer's Perspective." *Journal of Construction Engineering and Management*, 147(10), 04021118.

Kim, J., Youm, S., Shan, Y., and Kim, J. (2021). "Analysis of fire accident factors on construction sites using web crawling and deep learning approach." *Sustainability*, 13(21), 11694.

Kincelova, K., Boton, C., Blanchet, P., and Dagenais, C. (2020). "Fire safety in tall timber building: A BIM-based automated code-checking approach." *Buildings*, 10(7), 121.

Kines, P., Andersen, L. P., Spangenberg, S., Mikkelsen, K. L., Dyreborg, J., and Zohar, D. (2010). "Improving construction site safety through leader-based verbal safety communication." *Journal of safety research*, 41(5), 399-406.

Kobes, M., Helsloot, I., De Vries, B., and Post, J. G. (2010). "Building safety and human behaviour in fire: A literature review." *Fire Safety Journal*, 45(1), 1-11.

Koo, H. J., and O'Connor, J. T. (2022). "A Strategy for Building Design Quality Improvement through BIM Capability Analysis." *Journal of Construction Engineering and Management*, 148(8), 04022066.

Kretz, T. "The dynamic distance potential field in a situation with asymmetric bottleneck capacities." *Proc., International Conference on Cellular Automata*, Springer, 480-488.

Kretz, T., Bönisch, C., and Vortisch, P. (2010). "Comparison of various methods for the calculation of the distance potential field." *Pedestrian and evacuation dynamics 2008*, Springer, 335-346.

Ku, J.-H. (2019). "A study on evaluation plan of fire safety performance for public building." *Journal of the Korea Convergence Society*, 10(4), 25-32.

Kuligowski, E. D. (2016). "Human behavior in fire." *SFPE handbook of fire protection engineering*, Springer, 2070-2114.

Lai, S., Liu, K., He, S., and Zhao, J. (2016). "How to generate a good word embedding." *IEEE Intelligent Systems*, 31(6), 5-14.

Lebram, M., Backlund, P., Engström, H., and Johannesson, M. (2009). "Design and architecture of Sidh–a cave based firefighter training game." *Design and use of serious games*, Springer, 19-31.

Lee, S.-R. (2012). "An Experimental Study on the Fire Risk at Welding· Cutting Process." *Fire Science and Engineering*, 26(3), 60-66.

Li, H., Yang, X., Wang, F., Rose, T., Chan, G., and Dong, S. (2016). "Stochastic state sequence model to predict construction site safety states through Real-Time Location Systems." *Safety science*, 84, 78-87.

Li, H., Zhang, C., Song, S., Demirkesen, S., and Chang, R. (2020). "Improving tolerance control on modular construction project with 3D laser scanning and BIM: A case study of removable floodwall project." *Applied Sciences*, 10(23), 8680.

Li, N., and Guo, R.-Y. (2021). "Human behavior during emergency evacuation: Cell transmission model." *IEEE Access*, 9, 42463-42482.

Li, P.-J., Chen, C.-S., Chang, H.-P., Ho, H.-H., and Xie, B. (2020). "Explosion mechanism analysis during tunnel construction in the Zengwen Reservoir." *Tunnelling and Underground Space Technology*, 97, 103279.

Li, Y., Lu, X., Guan, H., Ying, M., and Yan, W. (2016). "A case study on a fire-induced collapse accident of a reinforced concrete frame-supported masonry structure." *Fire technology*, 52(3), 707-729.

Ling, W., Dyer, C., Black, A. W., and Trancoso, I. "Two/too simple adaptations of word2vec for syntax problems." *Proc., Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1299-1304.

Lopez-Aparicio, S., Grythe, H., Vogt, M., Pierce, M., and Vallejo, I. (2018). "Webcrawling and machine learning as a new approach for the spatial distribution of atmospheric emissions." *PloS one*, 13(7).

Lu, Y., Gong, P., Tang, Y., Sun, S., and Li, Q. (2021). "BIM-integrated construction safety risk assessment at the design stage of building projects." *Automation in Construction*, 124, 103553.

Ma, G., and Wu, Z. (2020). "BIM-based building fire emergency management: Combining building users' behavior decisions." *Automation in Construction*, 109, 102975.

Mahto, D. K., and Singh, L. "A dive into Web Scraper world." *Proc., 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, IEEE, 689-693.

Maluk, C., Woodrow, M., and Torero, J. L. (2017). "The potential of integrating fire safety in modern building design." *Fire Safety Journal*, 88, 104-112.

Mason, H., Karen Caruth, M., Lu, Q., and Cyr, D. R. (2019). "Review of Certification in Building a Successful Sonographer Program in China." *ADVANCED ULTRASOUND IN DIAGNOSIS AND THERAPY*, 3(3), 119-122.

Massimino, B. (2016). "Accessing online data: Web-crawling and information-scraping techniques to automate the assembly of research data." *Journal of Business Logistics*, 37(1), 34-42.

McDonald, M. A., Lipscomb, H. J., Bondy, J., and Glazner, J. (2009). "“Safety is everyone's job:” The key to safety on a large university construction site." *Journal of Safety Research*, 40(1), 53-61.

McInnes, L., Healy, J., and Melville, J. (2018). "Umap: Uniform manifold approximation and projection for dimension reduction." *arXiv preprint arXiv:1802.03426*.

Menczer, F., Fortunato, S., and Davis, C. A. (2020). *A First Course in Network Science*, Cambridge University Press.

Mohamed, S., Ali, T. H., and Tam, W. (2009). "National culture and safe work behaviour of construction workers in Pakistan." *Safety science*, 47(1), 29-35.

Mohan, S., and Zech, W. C. (2005). "Characteristics of worker accidents on NYSDOT construction projects." *Journal of Safety Research*, 36(4), 353-360.

Moon, S., Shin, Y., Hwang, B.-G., and Chi, S. (2018). "Document management system using text mining for information acquisition of international construction." *KSCE Journal of Civil Engineering*, 22(12), 4791-4798.

Morgan, J., Tietje, R., Wang, D., and Pattabhi, T. (2020). "Web Threat Investigation Using Advanced Web Crawling." Google Patents.

Mustika, N. R. (2018). "Automated Black Box Testing using Selenium Python." *International Journal of Computer Science and Software Engineering*, 7(9), 201-204.

Niwattanakul, S., Singthongchai, J., Naenudorn, E., and Wanapu, S. "Using of Jaccard coefficient for keywords similarity." *Proc., Proceedings of the international multiconference of engineers and computer scientists*, 380-384.

O'Toole, T. (2002). "Construction site safety roles Journal of Construction Engineering and Management 128."

Ombabi, A. H., Lazzez, O., Ouarda, W., and Alimi, A. M. "Deep learning framework based on Word2Vec and CNNfor users interests classification." *Proc., 2017 Sudan Conference on Computer Science and Information Technology (SCCSIT)*, IEEE, 1-7.

OSHA (2020). "Commonly Used Statistics." <https://www.osha.gov/data/commonstats>.

Ozturk, G. B. (2018). "Identifying the advantages of BIM in structural design." *Eurasian Journal of Civil Engineering and Architecture*, 2(2), 25-32.

Pajares, F., and Miller, M. D. (1994). "Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis." *Journal of educational psychology*, 86(2), 193.

Patel, D., and Jha, K. (2015). "Neural network model for the prediction of safe work behavior in construction projects." *Journal of Construction Engineering and Management*, 141(1), 04014066.

Paul, S., Mitra, A., and Dey, S. (2017). "Issues and challenges in web crawling for information extraction." *Bio-Inspired Computing for Information Retrieval Applications*, IGI Global, 93-121.

Pett, M. A., Lackey, N. R., and Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*, sage.

Pierce, B. (2013). "The Seasonal Timing of Work-Related Injuries October 2013." *Joint Statistical Meetings*, 2371-2381.

Pruskova, K., and Kaiser, J. "Implementation of BIM Technology into the design process using the scheme of BIM Execution Plan." *Proc., IOP Conference Series: Materials Science and Engineering*, IOP Publishing, 022019.

Qin, J., Liu, C., and Huang, Q. (2020). "Simulation on fire emergency evacuation in special subway station based on Pathfinder." *Case Studies in Thermal Engineering*, 21, 100677.

Rahmani, A., and Salem, M. (2020). "Simulation of fire in super high-rise hospitals using fire dynamics simulator (FDS)." *Electronic Journal of General Medicine*, 17(3), 5.

Rehman, S. U., Lee, S., Choi, J., and Kim, I. "ANALYSIS OF DEVELOPMENT OF OPEN BIM-BASED AUTOMATED RULE CHECKING SYSTEM IN KBIM PROJECT."

Rekabsaz, N., Lupu, M., and Hanbury, A. "Exploration of a threshold for similarity based on uncertainty in word embedding." *Proc., European Conference on Information Retrieval*, Springer, 396-409.

Rogsch, C., and Klingsch, W. (2012). "Basics of Software-Tools for Pedestrian Movement—Identification and Results." *Fire Technology*, 48(1), 105-125.

Rosseel, Y. (2012). "Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA)." *Journal of statistical software*, 48(2), 1-36.

Şahin, C., Rokne, J., and Alhajj, R. (2019). "Human behavior modeling for simulating evacuation of buildings during emergencies." *Physica A: Statistical Mechanics and its Applications*, 528, 121432.

Sandhu, R., Gill, H. K., and Sood, S. K. (2016). "Smart monitoring and controlling of Pandemic Influenza A (H1N1) using Social Network Analysis and cloud computing." *Journal of Computational Science*, 12, 11-22.

Shams Abadi, S. T., Moniri Tokmehdash, N., Hosny, A., and Nik-Bakht, M. (2021). "BIM-based co-simulation of fire and occupants' behavior for safe construction rehabilitation planning." *Fire*, 4(4), 67.

Shao, L., Zhang, H., Jia, M., and Wang, J. (2017). "Efficient and effective single-document summarizations and a word-embedding measurement of quality." *arXiv preprint arXiv:1710.00284*.

Shen, Y., Xu, M., Lin, Y., Cui, C., Shi, X., and Liu, Y. (2022). "Safety Risk Management of Prefabricated Building Construction Based on Ontology Technology in the BIM Environment." *Buildings*, 12(6), 765.

Shi, J., Dao, J., Jiang, L., and Pan, Z. (2019). "Research on IFC-and FDS-based information sharing for building fire safety analysis." *Advances in Civil Engineering*, 2019.

Shin, M., Lee, H.-S., Park, M., Moon, M., and Han, S. (2014). "A system dynamics approach for modeling construction workers' safety attitudes and behaviors." *Accident Analysis & Prevention*, 68, 95-105.

Sidani, A., Martins, J. P., and Soeiro, A. "BIM Health and Safety Support Framework for Construction." *Proc., DCE21-Symposium on Civil Engineering and Spatial Planning: Book of Abstracts*.

Sielicka, E., Choma, A., and Kowalczyk, D. "FIRE HAZARDS ON CONSTRUCTION AREAS." *OF PROFESSIONAL ACTIVITY AND CAREER WAY*, 21.

Sime, J. D. (2001). "An occupant response shelter escape time (ORSET) model." *Safety science*, 38(2), 109-125.

Smith, M., and Gorgoni, S. (2018). "An introduction to network analysis." *Networks of International Trade and Investment: Understanding globalisation through the lens of network analysis*, 1.

Su, Y., Mao, C., Jiang, R., Liu, G., and Wang, J. (2021). "Data-Driven Fire Safety Management at Building Construction Sites: Leveraging CNN." *Journal of Management in Engineering*, 37(2), 04020108.

Sun, Q., and Turkan, Y. (2019). "A BIM Based simulation framework for fire evacuation planning." *Advances in Informatics and Computing in Civil and Construction Engineering*, Springer, 431-438.

Sun, Q., and Turkan, Y. (2020). "A BIM-based simulation framework for fire safety management and investigation of the critical factors affecting human evacuation performance." *Advanced Engineering Informatics*, 44, 101093.

Sun, Q., Turkan, Y., and Fischer, E. C. (2022). "A building information modeling-fire dynamics simulation integrated framework for the simulation of passive fire protection in a mid-scale cross-laminated timber compartment: Numerical implementation and benchmarking." *Fire and Materials*.

Teo, A. L. E., Ofori, G., Tjandra, I. K., and Kim, H. (2016). "Design for safety: theoretical framework of the safety aspect of BIM system to determine the safety index." *Construction Economics and Building*, 16(4), 1-18.

Tran, S. V.-T., Nguyen, T. L., Chi, H.-L., Lee, D., and Park, C. (2022). "Generative planning for construction safety surveillance camera installation in 4D BIM environment." *Automation in construction*, 134, 104103.

Vijayarani, S., Ilamathi, M. J., and Nithya, M. (2015). "Preprocessing techniques for text mining-an overview." *International Journal of Computer Science & Communication Networks*, 5(1), 7-16.

Volk, R., Stengel, J., and Schultmann, F. (2014). "Building Information Modeling (BIM) for existing buildings—Literature review and future needs." *Automation in construction*, 38, 109-127.

Wang, N., Gao, Y., Li, C.-y., and Gai, W.-m. (2021). "Integrated agent-based simulation and evacuation risk-assessment model for underground building fire: A case study." *Journal of Building Engineering*, 40, 102609.

Wang, Z., and Jia, G. (2021). "A novel agent-based model for tsunami evacuation simulation and risk assessment." *Natural hazards*, 105(2), 2045-2071.

Wigglesworth, E. (2006). "Occupational injuries by hour of day and day of week: a 20-year study." *Australian and New Zealand journal of public health*, 30(6), 505-508.

Wolski, A., Dembsey, N. A., and Meacham, B. J. (2000). "Accommodating perceptions of risk in performance-based building fire safety code development." *Fire Safety Journal*, 34(3), 297-309.

Wright, S. (1960). "Path coefficients and path regressions: alternative or complementary concepts?" *Biometrics*, 16(2), 189-202.

Yang, H., Chew, D. A., Wu, W., Zhou, Z., and Li, Q. (2012). "Design and implementation of an identification system in construction site safety for proactive accident prevention." *Accident Analysis & Prevention*, 48, 193-203.

Yang, S., Wi, S., and Kim, S. (2018). "Development Methodology of Web Crawling Based on Physical Properties DB of Building Materials for the Efficiency of Building Energy Simulation." *한국생활환경학회지*, 25(4), 467-475.

Yang, X., Yang, X., and Wang, Q. (2020). "Pedestrian evacuation under guides in a multiple-exit room via the fuzzy logic method." *Communications in Nonlinear Science and Numerical Simulation*, 83, 105138.

Yin, Z., and Shen, Y. "On the dimensionality of word embedding." *Proc., Advances in Neural Information Processing Systems*, 887-898.

Yogana, E., and Latief, Y. "Development of system information of building code checking in planning and permitting phase to improve building code compliance based on work breakdown structure (WBS) using building information modeling (BIM)." *Proc.*, *Journal of Physics: Conference Series*, IOP Publishing, 012091.

Yoshioka, K., and Dozono, H. "The Classification of the Documents Based on Word Embedding and 2-layer Spherical Self Organizing Maps." *Proc., Proceedings of the 2019 11th International Conference on Machine Learning and Computing*, 357-361.

Zang, Y., Mei, Q., and Liu, S. (2021). "Evacuation simulation of a high-rise teaching building considering the influence of obstacles." *Simulation Modelling Practice and Theory*, 112, 102354.

Zhang, L., Wu, X., Liu, M., Liu, W., and Ashuri, B. (2019). "Discovering worst fire scenarios in subway stations: A simulation approach." *Automation in Construction*, 99, 183-196.

Zhang, Z., Huang, Y., Zhu, P., and Zhao, H. "Effective character-augmented word embedding for machine reading comprehension." *Proc., CCF International Conference on Natural Language Processing and Chinese Computing*, Springer, 27-39.

Zheng, C., He, G., and Peng, Z. (2015). "A Study of Web Information Extraction Technology Based on Beautiful Soup." *JCP*, 10(6), 381-387.

Zou, P. X., and Zhang, G. (2009). "Comparative study on the perception of construction safety risks in China and Australia." *Journal of construction engineering and management*, 135(7), 620-627.

APPENDICES

SOURCE CODE

1. Web crawling

```
Import requests
from bs4 import BeautifulSoup
from selenium import webdriver
import time
import random

random.seed(42)

#main sourcde url
#construction accident -> search keyword

if __name__ == "__main__" :
    m_url
="https://www.nytimes.com/search?dropmab=false&endDate=20200229&query=%22constructio
n%22%20accident&sort=best&startDate=20000101"

    session = requests.Session()
    req=session.get(m_url) #default is s_url
    soup = BeautifulSoup(req.text, 'html.parser')

    chrome_options = webdriver.ChromeOptions
    #chrome_options.add_argument("--incognito")
    driver =
webdriver.Chrome(executable_path='/Users/macbookpro/Project_construction/chromedriver')

    link = []
    driver.get(m_url)
    time.sleep(3)
```

```
for i in range(100) :
    a = driver.page_source
    b = BeautifulSoup(a, 'html.parser')
    #stream-panel > div.css-13mho3u > ol > li:nth-child(1) > div > div.css-1l4spti > a
    #stream-panel > div.css-13mho3u > ol > li > div > div.css-1l4spti > a
    #site-content > div > div:nth-child(2) > div.css-46b038 > ol > li:nth-child(1) > div > div > 
div > a
    c = b.select('#site-content > div >div>div> ol >li> div > div > div > a')


    for j in range(len(c)) :
        if c[j].attrs['href'] not in link :
            link.append(c[j].attrs['href'])
        else :
            pass


    driver.find_element_by_xpath('//*[@id="site-content"]/div/div[2]/div[2]/div/button').click()
    time.sleep(3)


#link copy
a_link = link.copy()
diff_link = []


#delete unsupported news
for i in range (len(a_link)) :
    print(a_link[i])
    if a_link[i][1] != '2':


        diff_link.append(a_link[i])
```

```
        else :
            continue

    for j in range(len(diff_link)):
        if diff_link[j] in a_link:
            a_link.remove(diff_link[j])
```

2. Collecting data from web

```
from bs4 import BeautifulSoup
import re

## This is main source for analyzing article

class ArticleInfo():

    def __init__(self, s_url) :

        session = requests.Session()
        req=session.get(s_url)
        self.soup = BeautifulSoup(req.text, 'html.parser')

    def article_body(self) :
        body = self.soup.find('div', {'class' : 'StoryBodyCompanionColumn'} )
        finalContent =''

        if body is not None :
            for itcontents in self.soup.findAll('p'):
                if itcontents == None :
```

```
                continue
            content = itcontents.getText()
            content = re.sub(r"\n+", "", content)
            finalContent +=content
        else :
            return False

        return finalContent
```

##reference: https://github.com/subhendusethi/nytimes-article-crawler/blob/master/crawler/nytimescrawler.py

```
    def article_title(self):
        titles = self.soup.select('title',{"data-rh":"true"})
        for title in titles:
            title = title.getText()
            break
        return title

    def article_date(self) :

        date = self.soup.find('li',{'class':'date'})
        if date != None :
            date = date.getText()

        else :
            date = self.soup.find('meta',{'name':'pdate'})
            if date !=None :
                date = date.get('content')
```

```
            if date !=None :
                date = str(date)
            else:
                none= "none"
                datelist.append(none)

        return date
'''
Get body, date, title from crawled data
'''
if __name__ == "__main__" :

    titles =[]
    bodies =[]
    datelist=[]
    lists = []

    for s_url in a_link :
        s_url = 'https://www.nytimes.com'+s_url
        ob = ArticleInfo(s_url)
        lists.append(s_url)

        body =ob.article_body()
        bodies.append(body)

        date = ob.article_date()
        datelist.append(date)
```

```
        title = ob.article_title()
        titles.append(title)
```

3. **Word2Vec Implementation**

```
From genism.models import Word2Vec
import networkx as nx
from matplotlib.backends.backend_pdf import PdfPages
# generate similar word with threshold and graph


key_list = ['fire','fell','collapse','building','people']
def network_gen(key_list,thres,sav_name) :
    set_of_g={}
    for key in key_list :


        sim_list = model.similar_by_vector(key, topn=100, restrict_vocab=None)
        network = [sim_list[i][0] for i in range(len(sim_list)) if sim_list[i][1] >= thres]
        network_df = [sim_list[i] for i in range(len(sim_list)) if sim_list[i][1] >= thres]

        set_of_g[key] = network


        df_word = pd.DataFrame(network_df, columns= ['Word', 'Similarity'])
        df_word.to_csv('Similar_words_100/'+key+str(thres)+'.csv')

    #generate graph
    #print(set_of_g)
    G = nx.Graph(set_of_g)
```

```
    nx.write_gml(G, "Network/"+sav_name+".gml")

    return "Success!"


if __name__ == "__main__" :

    model = Word2Vec(token_sentence, size = 400, window=5, min_count=200, workers=3, sg=0, 
iter=5)

    model.save("filename.model")

    '''

    check similarty between 5 keywords (total 10)

    fire - collapse, fire - fell, fire - people, fire - building

    collapse - fell, collapse - people , collapse - building

    fell - people , fell - building

    building - people

    '''

    sb1 = model.similarity('building','collapse')

    sb2 = model.similarity('building','fire')

    sb3 = model.similarity('building','people')

    sb4 = model.similarity('building','fell')

    sb5 = model.similarity('fire','collapse')

    sb6 = model.similarity('fire','fell')

    sb7 = model.similarity('fire','people')

    sb8 = model.similarity('collapse','fell')

    sb9 = model.similarity('collapse','people')

    sb10 = model.similarity('people','fell')

    '''

    print('building and collapse', model.similarity('building','collapse'))  ## similarity between 
words
```

```
print('building and fire', model.similarity('building','fire'))
print('building and people', model.similarity('building','people'))
print('building and fell', model.similarity('building','fell'))
print('fire and collapse', model.similarity('fire','collapse'))
print('fire and fell', model.similarity('fire','fell'))
print('fire and people', model.similarity('fire','people'))
print('collapse and fell', model.similarity('collapse','fell'))
print('collpase and people', model.similarity('collapse','people'))
print('people and fell', model.similarity('people','fell'))
'''
sim_between = [('building and collapse' , "%.3f" % sb1),('building and fire',"%.3f" % sb2),
    ('building and people ',"%.3f" % sb3),
    ('building and fell',"%.3f" % sb4),('collapse and fell', "%.3f" % sb8),
    ('collapse and people',"%.3f" % sb9),('fire and collapse',"%.3f" % sb5),
    ('fire and fell', "%.3f" % sb6), ('fire and people',"%.3f" % sb7),
    ('people and fell', "%.3f" % sb10)]
df_sim_between = pd.DataFrame(sim_between, columns=['Words', 'Similarity'])
df_sim_between.to_csv('Figure/sim_between_0521.csv')

# top 20 data for graph coloring, top 50 data for showing, top 100 data for make network graph

sim_list_fire = model.similar_by_vector('fire', topn=20, restrict_vocab=None)
sim_list_fall = model.similar_by_vector('fell', topn=20, restrict_vocab=None)
sim_list_collapse = model.similar_by_vector('collapse', topn=20, restrict_vocab=None)
sim_list_building = model.similar_by_vector('building', topn=20, restrict_vocab=None)
sim_list_people = model.similar_by_vector('people', topn=20, restrict_vocab= None)
```

```
# list of similiar 20 words -> generating color in UMAP

sim_fire_word = []

for i in range(len(sim_list_fire)) :

    sim_fire_word.append(sim_list_fire[i][0])

#print(sim_fire_word)

sim_fell_word = []

for i in range(len(sim_list_fall)) :

    sim_fell_word.append(sim_list_fall[i][0])

#print(sim_fall_word)

sim_collapse_word =[]

for i in range(len(sim_list_collapse)):

    sim_collapse_word.append(sim_list_collapse[i][0])

#print(sim_collapse_word)

sim_building_word = []

for i in range(len(sim_list_building)):

    sim_building_word.append(sim_list_building[i][0])

#print(sim_building_word)

sim_people_word = []

for i in range(len(sim_list_people)):

    sim_people_word.append(sim_list_people[i][0])

#print(sim_people_word)



keywords = ['fire','fell','collapse','building','people']

keyword_dict = {}
```

```
    for keyword in keywords :
        a = 'sim_%s_word'%(keyword)
        ##local varible same name with a

        keyword_dict[keyword] = locals()[a]

    df_keyword = pd.DataFrame.from_dict(keyword_dict)
    #df_keyword.to_csv('keyword_simlist.csv')

    fig, ax = plt.subplots(figsize=(12,4))
    ax.axis('tight')
    ax.axis('off')
    the_table = ax.table(cellText = df_keyword.values, colLabels = df_keyword.columns, loc = 'center')
    pp = PdfPages("simlist.pdf")
    pp.savefig(fig,bbox_inches = 'tight')
```

4. UMAP Implementation

```
from umap import UMAP
from gensim.models import Word2Vec
'''
From word2vec.py...
model, sim_*_word
'''
if __name__ == "__main__" :
    #UMAP
```

```
model = Word2Vec.model.load("path")


reducer = UMAP(n_neighbors =5, min_dist =0.1, n_components = 2, verbose = True)

X = model[model.wv.vocab]

#list of word

X_l = list(model.wv.vocab)


#Embedding to 2 dimension

cluster_embedding = reducer.fit_transform(X)

#get coordination

df = pd.DataFrame(cluster_embedding)


fig = plt.figure()

fig.set_size_inches(50,30)

ax = fig.add_subplot(1,1,1)


ax.scatter(df[0],df[1])


for i, txt in enumerate(X_l):


    if txt == 'fire' :

        ax.annotate(txt, (df.loc[i][0],df.loc[i][1]), color = 'red', fontsize=30)

    elif txt in sim_fire_word :

        ax.annotate(txt, (df.loc[i][0],df.loc[i][1]), color ='red', fontsize=8)


    elif txt =='fell' :
```

```
        ax.annotate(txt, (df.loc[i][0],df.loc[i][1]),color='green', fontsize=30)
    elif txt in sim_fell_word :
        ax.annotate(txt, (df.loc[i][0],df.loc[i][1]), color ='green', fontsize=8)
    elif txt =='collapse' :
        ax.annotate(txt, (df.loc[i][0],df.loc[i][1]), color = 'blue', fontsize=30)
    elif txt in sim_collapse_word :
        ax.annotate(txt, (df.loc[i][0],df.loc[i][1]), color ='blue', fontsize=8)
    elif txt =='building' :
        ax.annotate(txt, (df.loc[i][0],df.loc[i][1]), color ='orange', fontsize=30)
    elif txt in sim_building_word :
        ax.annotate(txt, (df.loc[i][0],df.loc[i][1]), color ='orange', fontsize=8)
    elif txt =='people' :
        ax.annotate(txt, (df.loc[i][0],df.loc[i][1]),color = 'purple', fontsize=30)
    elif txt in sim_people_word :
        ax.annotate(txt, (df.loc[i][0],df.loc[i][1]), color ='purple', fontsize=8)
    else :
        ax.annotate(txt, (df.loc[i][0],df.loc[i][1]), fontsize=8)


plt.title ("Word Embdding results with UMAP", fontsize = 20)
plt.savefig('fire_word2vec_0521_best_fell.pdf')
```

VITA

JAEHONG KIM

Candidate for the Degree of Doctor of Philosophy

Dissertation: A MULTIMODAL APPROACH TO IMPROVE FIRE SAFETY ON CONSTRUCTION SITES

Major Field: Civil Engineering

Biographical:

Education:
Ph.D. Civil and Environmental Engineering
Oklahoma State University, Stillwater, Oklahoma, US.
Dec. 2022

M.S., Architectural and Civil Engineering
KYUNGPOOK NATIONAL UNIVERSITY, Daegu, South Korea
Feb. 2012

B.S., Architectural Engineering
KYUNGPOOK NATIONAL UNIVERSITY, Daegu, South Korea
Feb. 2010

Experience:
Graduate Research Assistant and Graduate Teaching Assistant
Oklahoma State University, Civil and Environmental Engineering
Aug. 2017 – Dec. 2022

Researcher
Digital Space Research Lab., Virtual Builders Co., Ltd., Seoul, South Korea
Feb. 2012 – Mar. 2015

Research Assistant
Kyungpook National University, Daegu, South Korea
Mar. 2008 – Feb. 2012