Workplace Hazard Identification through Near-Miss Reports: A Social Network Analysis

D P Anggraini[1], R Ambarwati[*2], Dedy[3], M T Alimova[4]

[1,2]Department of Management, University Muhammadiyah Sidoarjo, Indonesia

[3]Health Safety Environment, PLN Nusantara Power, Surabaya, Indonesia

[4]Department of Management, Samarkand Institute of Economics and Service, Uzbekistan

E-mail: ditaputrianggraini001@gmail.com[1], ritaambarwati@umsida.ac.id[2], dedychemist@gmail.com[3], alimovamashhura@sies.uz

Abstract. Occupational Safety and Health (OHS) aims to create a safe and healthy working environment. The company engaged in power generation and experienced many near-miss cases from 2016 to 2023. Near-miss incidents are an early signal of the potential for severe accidents in the future if not resolved immediately. Identifying and addressing potential hazards early is necessary to prevent fatal work accidents. This study aims to analyze the reporting pattern of near-miss incidents by company employees through the IZAT application with a social network analysis (SNA) approach. This method identifies relationships and interaction patterns between employees in reporting near-miss incidents, thus revealing objects that need special attention. Data was obtained from September 2020 to July 2023. The analysis showed that 276 employees actively reported various near-miss incidents through IZAT (Zero Accident Assistant Application). This pattern reflects employees' sensitivity to potential hazards in their work environment. The findings of this study make a positive contribution to the improvement of the occupational safety and health management system in the company regarding the mitigation of potential hazards based on incoming report data. The recommendations are related to occupational safety aspects that must be prioritized for management improvement. The implication of this research is to increase employee awareness in identifying and reporting potential hazards to prevent work accidents in the future

Keywords: safety culture; near miss; occupational health and safety; reporting; social network analysis

1. Introduction

The application of Occupational Safety and Health (K3) is the key to Occupational Health (K3) is the key to creating a safe work environment [1]. Occupational Safety and Health is an effort to ensure a safe and healthy work environment to prevent disability accidents, accidents, and even death caused by work accidents. It is essential to protect workers and safeguard an organization's human resources[2]. Work safety in all workplaces, on land, in water, in the ground, underground, and in the air, has been regulated in Article 2 Paragraph 1 of Law of the Republic of Indonesia Number 1 of 1970 concerning Work Safety. However, in practice, the implementation of the law has yet to be fully optimal. There is a high rate of workplace accidents in several business sectors. The company, which has been operating power plants since 1995, is trusted to manage more than 40 generating units throughout Indonesia. This company has a high level of risk of work accidents because its workers directly interact with electrical components that have low and high voltage. Based on accident statistics in the company from 2016 to 2023, there was one serious injury incident in 2016 and no similar incidents in subsequent years. Meanwhile, there was a fatal incident, *with* details of 1 case in 2016, 1 case in 2017, 4 cases in 2018, 1 case in 2019, and 2 cases in 2020. Meanwhile, there were no events fatality for the following years (2021-2023).

According to Heinrich, in his research, unsafe actions cause 88% of accidents that occur in the work environment, unsafe environmental conditions cause 10%, and God's destiny causes 2%. Heinrich's accident pyramid states that for every fatal accident, there are 30 significant accidents or serious injuries that can cause damage to equipment, assets, and property, as well as 300 minor accidents, as well as 600 near-misses, and exposure to 30,000 hazards due to unsafe actions and unsafe conditions. This pyramid of work accidents illustrates that to prevent fatal accidents in the workplace, efforts need to be made to reduce the incidence of near harm in the work environment. Thus, the chances of fatal accidents and other incidents occurring before reaching a fatal accident can be reduced [3]. Problems related to work accidents in the company encourage companies to manage K3 risks by innovating to design the IZAT Application (Zero Accident Assistant Application) 2.0. This application is designed to facilitate the improvement of the efficiency of all business processes in the aspect of Occupational Safety and Health (K3) by encouraging the development of a culture of reporting and corrective action on the findings of patrol results. The IZAT application (Zero Accident Assistant Application) 2.0 has a status search feature; reporters can categorize dangerous conditions into four types: positive (non-hazardous conditions), near misses, unsafe actions, and unsafe conditions.

Near-miss reporting has a vital role to play in preventing accidents in the workplace. By reporting near misses, companies can identify potential hazards and risks before a fatal accident occurs. An in-depth analysis of the report allows the identification of root causes and early indicators to minimize the risk of light, severe, and even fatal accidents. The ISO 45001 standard defines a near miss as a work-related incident with no injury or health impairment but the potential to cause it. A near miss is not an accident but a latent danger signal that requires corrective action. This is the urgency of near-miss reporting involving all permanent, non-permanent, and skilled workerso support the company's zero accident vision [4]. Research conducted by Nikhil in his research entitled A system dynamics model for near-miss reporting in complex systems [5] examines the role of near miss reports in improving safety in organizations managing highly safe complex systems. Using System Dynamics methodology, this research highlights the need to consider employee interactions and management decisions that influence employee behavior. These findings illustrate the trade-off between near miss reporting and hours worked, thereby assisting management in addressing worker fatigue relevant to highly safe complex systems. Seok Ki Lee investigates near-miss reporting in Korean manufacturing workplaces, focusing on the relationship between such activities, safety culture, and industrial accidents. This quantitative study finds significant perception differences regarding near-miss events based on whether near-miss reporting activities are conducted [6]. Yi Yang's cross-sectional survey collected data from 920 registered nurses across eight tertiary hospitals in China to explore the relationship between patient safety culture and near-death incident reporting. The study found a positive correlation between patient safety culture and nurses' intention to report near-death incidents, with the severity of incidents potentially moderating this relationship [7].

As explained above, the importance of near misses has been recognized by many researchers. However, most previous studies have focused more on individual factors and organizational culture, which are less effective in capturing the complexity of social interactions and relationship patterns. Few researchers have analyzed the relationships between entities in a social network. Social Network Analysis methods can provide new perspectives by exploring social interactions and network dynamics, resulting in simple network visualizations. In addition, most previous studies have been conducted in construction companies or hospitals. Therefore, this study aims to fill the gap in the literature related to risk reporting in the power generation industry. This study analyzes text mining through the IZAT (*Zero Accident Assistant Application*) 2.0 application on reporter responses regarding near miss risk findings in the Power Generation Unit. The purpose of this study is to identify the tendency of reporter responses to the near-miss category, as well as to increase employee awareness in identifying potential hazards. The theoretical contribution is to explain the relationship between factors that influence near miss reporting behavior through network visualization. Practically, this study provides suggestions and evaluations to improve and enhance the K3 culture in power generation companies, which can be applied universally.

2. Literature Review

2.1 Previous Research

Previous studies have examined various aspects related to workplace safety, including safety culture, safety risks, and safety reporting. These studies used a variety of research methods to explore and analyze factors that influence safety in the workplace. The following is a summary of some relevant studies, categorized by the main research themes: safety culture, safety risks, and safety reporting.

Table 1. Summary of Relevant Empirical Studies on Occupational Safety and Health

Theme	Reference	Research Focus	Method
Safety Culture	Tracy et al. (2020)	Risk perception and safety culture as management tools	Risk assessment and analysis
	Minh et al. (2019)	Interactive effects of safety culture on construction project	SEM-PLS
	Yingbin et al. (2019)	Factors driving safety culture	SEM-PLS
	Emily et al. (2020)	Analysis of occupational health and safety	Data analysis with leximancer
	Jiangshi et al. (2020)	Characteristics of safety culture deficiencies	Why because analysis (wba)
	Mambwe et al. (2021)	Ohs improvement management strategy	Quantitative research
	A Jalil et al. (2021)	Construction safety culture	Questionnaire with numerical scale
Safety Risk	Xingwei et al. (2019)	Development of offshore major hazard risk prevention	Spar-h and petro-hra
	Sooyoung et al. (2020)	Safety risk generation and control model	SLR
	L Ding et al. (2020)	Generic bow tie framework model to prevent occupational accident risks	Case study
	Nasim et al. (2020)	Applying case studies to risk identification	Case study
	Yuncan et al. (2019)	Use of software for safety risk prevention	Bayesian network analysis
	V Dat et al. (2020)	Practical implications of risk theory in the covid-19 situation	CFA model

Theme	Reference	Research Focus	Method
Safety Reporting	Davood et al. (2020)	Statistical analysis of perceived safety risks to construction workers	Statistical analysis
	Kabul et al. (2022)	Identify potential environmental hazards	Qualitative
	Tashia et al. (2023)	Evaluation of routine reporting program	Qualitative and interviews
	Mayangkara et al. (2021)	Implementation of a system for recording and reporting occupational accident	Qualitative
	Soltanzadeh et al. (2019)	Analysis of accident severity in the chemical industry	Retrospective descriptive-analytic
	Derakhshan et al. (2021)	Analysis of unsafe condition factor reports and unsafe acts and identification of accident types	Retrospective descriptive-analytic
	Hasanspahic et al. (2020)	Analysis of crew willingness to report near miss events	Qualitative research methods

Table 1 describes previous research findings related to safety culture, risk safety and safety reporting. Safety culture is one of the important factors that influence occupational safety. Recent studies have provided valuable insights into the influence of safety culture on safety behaviour, motivation and the effectiveness of safety initiatives. In healthcare, research by Reis et al. showed that a strong safety culture can improve teamwork and patient safety. Research by Al-Bayati showed that construction safety culture and safety environment can influence safety behaviour and motivation [8]. In construction, research by Mambwe et al. showed that safety culture is a critical component for achieving successful outcomes in implementing occupational health and safety management strategies [9]. Overall, these studies show that safety culture is an important factor to consider in improving occupational safety.

Several studies have shown that implementing a safety culture provides a number of significant benefits. Reis et al's analysis of the dimensions of the Hospital Survey on Patient Safety Culture led to an understanding of the positive impact safety culture has on cooperation and safety in the medical field. The assessment also has the potential to improve co-operation and learning in healthcare environments. Meanwhile, Al-Bayati's research examined how the environment and safety culture can influence safety behaviour and motivation in the construction industry, especially regarding safety outcomes. Meanwhile, Tetzlaff et al's research conducted a retrospective evaluation of health and safety reports in the mining industry, which provided deep insights into the benefits of safety culture in the industry [10]. All three studies highlight the benefits of safety reporting and culture across various industries, from mining to construction to healthcare.

How safety attitudes influence safety risk perception and hazard identification is a new subject of safety risk research [11]. One of them is research on how safety attitudes influence risk perception and hazard identification. The research shows that the disposition of individuals in the workplace can influence their perception of safety and hazards. Another study provided insights into safety risk management at the manufacturing stage, especially in the automotive supply industry [12]. In addition, Lbahar et al introduced a new method for occupational health and safety risk assessment by applying the Pythagorean Fuzzy AHP process and a fuzzy inference system. The methodology enables more comprehensive and flexible risk assessment by incorporating complexity in decision-making. With such advanced techniques, the researchers advance risk analysis methodologies in the field of occupational health and safety and offer innovative methods for risk management.

2.2 Social Network Analysis (SNA)

Social Network Analysis is a science that studies explicitly human relationships and relationships using graph theory [13]. The illustration of social network analysis can be seen through its presentation of a network that relies on two main components for disseminating information on communication networks: individuals (nodes) and connections (edges). Social network analysis is a method to quantitatively analyse groups' structure, beliefs, and evolutionary processes by modelling the relationships between groups as nodes and links. This will provide knowledge to understand the relationship structure by describing the relationship between them as an edge [14]. Social network analysis assumes that network members are related to each other and that their behaviour is largely influenced by the pattern of relationships reflected in the network structure.

3. Research Method

This study uses the text mining method with Social Network Analysis (SNA) to analyze the risk findings of the near miss category between September 2020 and July 2023. Text mining extracts significant information and context from unstructured text-based data [15]. Meanwhile, SNA is used to identify critical actors in the dissemination of information and visualize relationships between individuals as nodes and bonds [16]. This study aims to map social networks based on incident findings that are reported to provide deep insights into the factors that influence the dissemination of information and risks in the Company's work environment. The concept of the research flow can be seen in Figure 1.

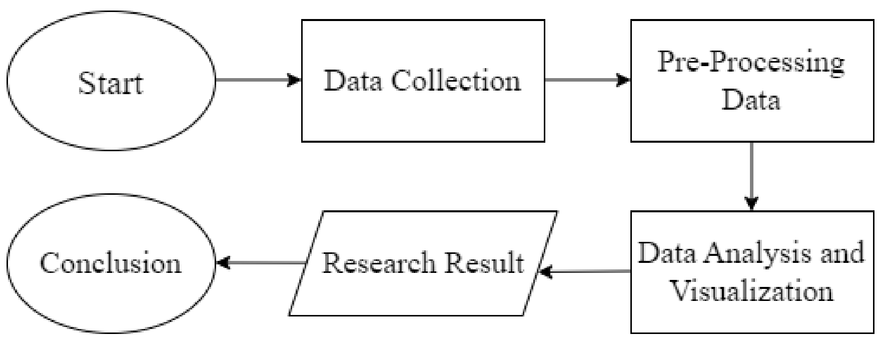

Figure 1. The Concept of The Research

The initial stage in the research process is the collection of historical data through the IZAT application (Zero Accident Assistant Application). The data is then extracted (scrapping) in Excel form, including the scope of three contents: near-miss findings, finding titles, and generating units. Data pre-processing is required before entering the analysis and visualization stage. The purpose of pre-processing in this study is to ensure that the data collected in text form consists only of those terms that have relevance to be analyzed. This process involves four main stages: case folding, tokenization, normalization, and filtering. As an analytical tool to identify the structure of the formed property network and visualize it, Gephi 0.9.2 software is used. Gephi is an interactive tool used for visualization and examination or an assessment platform for various types of simple and complex networks and dynamic and hierarchical graphs [17]. In realizing the network visualization results, we use the ForceAtlas2 methodology as the algorithm graph layout to understand the layout procedure d the impact of various parameters [18].

3.1. Data Collection

In 2020, the Head Office HSSE Division created the Zero Accident Assistant (IZAT 2.0) application equipped with a K3 nonconformity report feature. The IZAT application is planned to support scheduled patrol activities involving all workers in power generation companies. Data from September 2020 to July

2023 showed 1,458 near-miss reports from 36 power plants. The data collection process uses the jupyter notebook scrapping tool with a Python programming language, and the " title " content is formulated in a query. The successfully extracted data is saved in CSV format to enter the next stage of processing [19]. The data collection results were identified as the primary data sources in this study, presented in Table 2.

Table 2. Data Withdrawal Results

	2020	**2021**	**2022**	**2023**	**Total Reporting**
Number of Findings	78	621	431	328	1,458

Table 2 shows the number of near-miss findings each year in the company. In 2020, the number of reports was still small because whistleblowers needed to adapt to new processes through the newly launched IZAT App in September and the COVID-19 pandemic. In 2021, there was a drastic spike in near-miss findings from 78 to 621 cases; this happened because more and more generating units joined. However, from 2022 to 2023, there has been a decrease in near-miss findings, indicating increased awareness and action on near-miss findings.

3.2. Pre-processing Data

Preprocessing is done to clean up data sources containing unnecessary words. Data preprocessing aims to eliminate noise in report data, such as abbreviations and non-standard words that are difficult for computers to translate [20]. Preprocessing consists of several stages:

a. Case folding is a stage that aims to convert all letters in the title into lowercase letters, so that there are no capital letters left in the title. The purpose of this process is to make the characters in the headline uniform [21].
b. Tokenising can take the form of characters, words, or sub-words. Thus, the tokenisation process can be broadly classified into three types, namely character-level, word-level, and sub-word-level tokenization [22].
c. Data normalisation is carried out to change words that are not standard or according to the correct spelling. At this stage it is carried out using a dictionary database of standard and non-standard language words that is created by yourself based on the reporter's response data [23].
d. Filtering can be called stopword removal, which is to remove words that are not important in the classification and reasoning process [24]. After that, the data is saved in csv form.

After going through the filtering stage in the preprocessing process, the data is then processed using Wordij. The processed results are in data formats such as Excel, NET, STP, and STW. Wordij is a text analysis tool that counts the frequency of each word and word pair occurrences in a given slice window (w) of text [25]. The results from Wordij will be classified according to predetermined categories based on the research problem.

3.3. Data Analysis and Visualization

The normalized information was processed using Wordij to extract the most important aspects of the dataset to be analyzed. After data preprocessing, social network modeling was performed using the social network analysis (SNA) method. Network visualization is done with the Gephi 0.9.2 application so that nodes representing users and edges representing relationships between actors in the social network can be seen. This tool aims to help find patterns and distill dynamic, iterative visualizations in making hypotheses.

4. Result and Discussion

The data source was obtained through organizational records obtained from the power generation unit through the IZAT application, and the initial data was previously 1.458. After data processing, using the jupyter notebook scrapping tool with the Python programming language to reduce data duplication, the amount of data was changed to 1.360. Of these, 98 data were duplicates. The next stage involved pre-processing in query-based reporting data collection. The aim is to improve data quality, remove noise, and clarify the information contained in the data.

4.1. Case Folding

Table 3. Stages of Case Folding

Before	After
CCR hallway ceiling roof almost collapsed	ccr hallway ceiling roof almost collapsed
January fire extinguisher check has not been carried out	january fire extinguisher check has not been carried out
Oil leakage from conveyor motor	oil leakage from conveyor motor
Apar tube no marking signs	apar tube no marking signs
Turbine area evacuation marker (EXIT) damaged	turbine area evacuation marker (exit) damaged

In the case folding stage, as shown in Table 3, in addition to converting capital letters into lowercase letters, the case folding process also removes punctuation marks and characters other than letters. In news headlines, punctuation marks and characters other than letters are considered as delimiters. Removing punctuation and characters other than letters aims to simplify the text so that it is easier to process in the next stage. For example, the phrase ‘CCR hallway ceiling roof almost collapsed’, the capital letters in the words ‘CCR’ are changed to lowercase letters so that it becomes ‘ccr hallway ceiling roof almost collapsed’ and the phrase ‘January fire extinguisher check has not been carried out, the capital letters in the words ‘January’ are changed to lowercase letters so that it becomes ‘january fire extinguisher check has not been carried out’.

4.2. Tokenizing

Table 4. Stages of Tokenization

Raw data	Preprocessed data
ccr hallway ceiling roof almost collapsed	'roof', 'ceiling', 'hallway', 'ccr', 'collapse'
january fire extinguisher check has not been carried out	'checking', 'fire extinguisher', 'january', 'not yet', 'done'
oil leakage from conveyor motor	'leak', 'oil', 'motor', 'conveyor'
apar tube no marking signs	'tube', 'fire extinguisher', 'sign', 'marking'
turbine area evacuation marker (exit) damaged	'marker', 'evacuation', 'exit', 'area', 'turbine', 'damaged'

Table 4 illustrates the tokenisation stage process. At this stage, the text is separated into meaningful units, such as words or short phrases [26]. Tokenization can be done automatically by using software called Scrape Library. The process of filtering and extracting keywords is essential to get genuinely relevant words and represent the main topic of the whole data. The keyword filtering and extraction process is done using the frequency method. The words that appear most often in the text will be considered as keywords. The results of tokenization are pieces of words that are then filtered and extracted as the primary keywords from the entire text data. The extracted keywords then form the nodes in the network mapping for social network analysis (SNA) visualization.

4.3. Filtering

Table 5. Filtering Stage

Before	After
ccr hallway ceiling roof almost collapsed	hallway ceiling roof collapsed
january fire extinguisher check not done	january fire extinguisher check done
oil leakage from conveyor motor	conveyor motor oil leakage
fire extinguisher tube no marking sign	fire extinguisher tube marking sign
turbine area evacuation (exit) markers damaged	turbine area exit evacuation markers damaged

In the filtering stage, as shown in Table 5, stopwords are removed from the text data. Stopwords are words that are not unique or characteristic a document or do not convey any significant message in the text or sentence [27]. Removing stopwords can reduce the amount of text data that needs analysis. In addition to stopwords, several other things can be omitted from text data during the filtering process, such as typos or meaningless words. Typos can interfere with the analysis results, as they can cause the analysis to produce patterns or trends that do not exist. Filtering aims to filter and retrieve only those essential words in the text.

4.4. Pre-processing data

Table 6. Pre-processing Result

Content Focus	Year	Word Count	Unique Word	Frequency
‘Title’ fiture	2020-2023	2.306	276	8.36

Pre-processing of the data in Table 6 is an integral part of social network analysis (SNA). Text preprocessing is the initial stage in text mining, which produces a set of term indexes that can represent documents. This stage converts the raw data into a format that is ready for analysis, thereby increasing the accuracy of social network analysis (SNA) results [28]. Attributes such as total word count, unique words, and word frequency inform the textual characteristics of near miss incident content. The total word count reflects how often the word appears in the incident report in the IZAT application. Unique words show the diversity of vocabulary in the text. The more unique the words, the more variations there are in the text. Word frequency shows the intensity with which certain words appear throughout the text. These three attributes are useful for understanding the relationship patterns between words and between words and the whole text in social network analysis (SNA). The data shows that near miss content has 2,306 words, 276 unique words, and 8.36 average words.

Table 7. Identification of Thematic

Place		**Response**		**Cause**		**Tools**		**Condition**	
area	0,030%	patrol	0,043%	pintu (door)	0,019 %	lantai (floor)	0,015 %	rusak (broken)	0,019 %
tempat (place)	0,025%	pakai (use)	0,020%	lampu (light)	0,014 %	sampah (rubbish)	0,015 %	bekas (used)	0,013 %
ruang (space)	0,014%	patroli (patrole)	0,011%	air (water)	0,011 %	kabel (cable)	0,012 %	mati (die)	0,009 %
unit	0,012%	wtp	0,006%	rambu (signs)	0,006 %	hydrant	0,009 %	jalan (street)	0,008 %
ccr	0,009%	cuci (wash)	0,005%	tangga (stairs)	0,005 %	barang (item)	0,009 %	berserakan (scattered)	0,007 %
ruangan (room)	0,007%	rutin (routine)	0,005%	limbah (waste)	0,004 %	apar (appear)	0,008 %	lepas (retrieved)	0,007 %
gedung (bulidin g)	0,007%	mandi (bathing)	0,004%	oil	0,004 %	box	0,008 %	sesuai (suitable)	0,006 %
penemp atan (placem ent)	0,006%	temuan (finding)	0,004%	dinding (wall)	0,003 %	papan (board)	0,007 %	habis (discharge d)	0,005 %
pos	0,005%	kerja (work)	0,004%	fire	0,003 %	ac	0,006 %	safety	0,005 %
kamar (room)	0,005%	penerang (lighting)	0,004%	water	0,003 %	alat (tool)	0,006 %	bocor (leak)	0,004 %

Table 7 shows the top 10 thematic words for each topic. It can be seen that 'place' in Topic 1 and "response" (Topic 2) are the two topics that reporters are most concerned about, with a particular focus on "patrol" (0.043%), "area" (0.030%), and "place" (0.025%). The word patrol achieved the highest value, which illustrates that most of the reports of potential hazards came from routine patrol activities in the work area. The prominent use of the words area and place also indicates that there are many potential hazards associated with specific locations in the work area. Therefore, this stage provides a visual representation of the most frequent words in the report, which can be used to analyse the root causes of potential hazards in the work environment.

Social network analysis (SNA) visualization uses nodes and edges to show the network structure graphically [29]. Nodes are represented as points or nodes, while edges are represented as lines or relationships between nodes. To include a node in the network, the criteria used are relevance to the topic being studied, data availability and participation in the network [30]. Meanwhile, to include an edge in the network, the criteria used are the existence of a real relationship between nodes and the measurability of the relationship. With this visualization, it can be clearly seen how the nodes in the network are connected through the edges that describe the relationship between these entities. The number of occurrences of a word or node connected to an edge is called the degree [31]. This degree indicates the total edges connected to a specific node in the graph network structure. The results of the property network are presented in Table 8.

Table 8. Network Property Result

Network Property	Value
Nodes	276
Edges	111
Average Degree	0,402
Average Weight Degree	1,783
Network Diameter	5
Modularity	0,799
Average Path Length	2,201

Table 8 shows the results of calculations on organizational structure network properties. Nodes are representations of actors (users) in a social network. If more than one node is connected to another node, the network can bquite active with many interacting actors [32]. The organizational structure network studied in this research has 276 nodes with 111 edges. The number of edges in a social network is less than that of nodes, which means that not all nodes in the network are connected to one node and another node but still form a network. Average degree shows the average number of connections a node has with others [33]. The higher the average degree value, it means that the greater the number of edges connecting the nodes, the faster the information dissemination will be. The average degree value in this organizational structure network is 0.402, which indicates that the dissemination of information in this organizational structure is very weak. Diameter is the greatest distance between 2 nodes [34]. In an organizational structure network, the network diameter is 5. The smaller the network diameter, the faster the spread of information between actors. Modularity describes the extent to which the network is separated into other groups within the network [35]. The higher the modularity value, the more clearly a network will be formed. It can be interpreted that each network obtained forms a different community where this community makes a network have more specifications for the community. The modularity value in the organizational structure network is 0.799. Average path length shows the average distance between a node and other nodes. The smaller the average path length value,he faster the information spreads [36]. The shows the number 2.201, which means that on average a node when connected to another node must pass through 2 nodes first so that the distribution of information in this network is fast [37].

After calculating network properties, the next step is creating a network model. Creating a network model visualization using Gephi software using an undirected graph type without paying attention to the direction of the relationship at a node, namely indegree (destination node) and outdegree (origin node). This research uses the ForceAtlas2 algorithm, which is a spatial layout algorithm for a web network. ForceAtlas2 can group nodes in a community, making it easier to observe and analyze.

4.5. Visualization Result

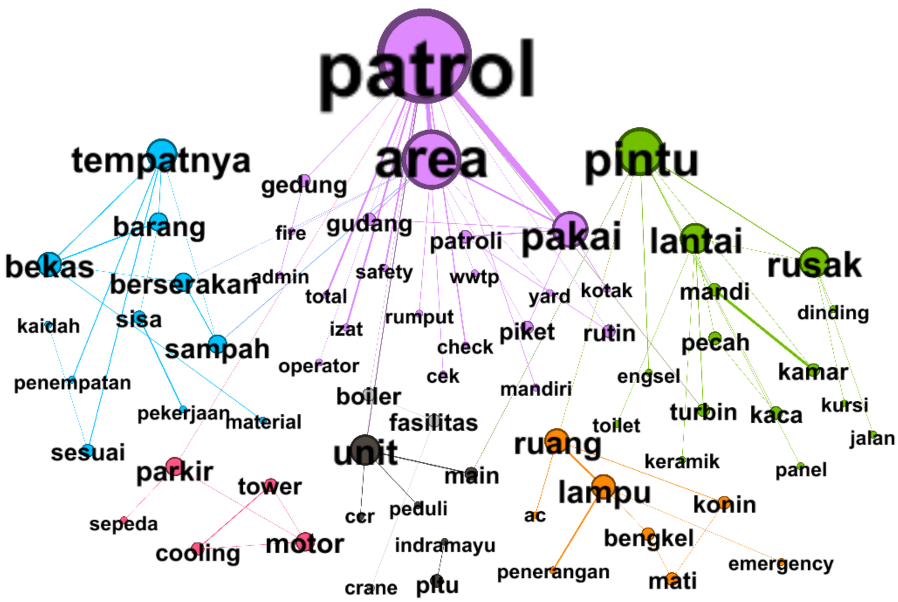

Figure 2. Network Visualization Map

The visualization results of near miss findings in Figure 2 show the relationship between nodes and edges that are part of the social network in the IZAT application. Cluster refers to a group of closely connected nodes in a network. From the results of the social network visualization, six clusters are closely interconnected: purple, green, blue, orange, pink, and black. Purple clusters with thick edges between nodes indicate stronger connection or interaction between these nodes. This shows that the findings dominate the near miss incident reports submitted by the reporter during routine patrols in the work area. Reporting often uses the word patrol and findings of potential dangers tend to relate to specific location specifications such as warehouses, buildings and other areas. Furthermore, the green cluster reflects the use of public facilities such as doors, glass, hinges and so on. This finding indicates damage to public facilities in the power generation industry. Blue clusters indicate the causes and conditions of the report findings. Reporters often use the words goods, used, rubbish and so on. The orange cluster often talks about aspects of the place and condition of the report's findings. The reporter often talks about space, lights, workshops, lighting, blackout, etc. Then, the pink cluster often discusses aspects of where the report was found. The reporter often uses the words parking, tower and others. However, this node is not strong enough because only one node is interconnected. Meanwhile, the gray cluster often uses the word unit, indicating that the pattern of reporting potential hazards is focused on certain work locations or units.

Table 7 displays five thematic word categories that reflect the dominant word patterns and findings in reported near miss incidents. Meanwhile, the six clusters in the social network analysis visualization show the main categories and interactions between factors causing incidents. Both have a correlation with each other. The purple cluster is related to the response category. The green cluster is related to the tools category. Furthermore, the blue cluster is related to the cause and condition category. Meanwhile, the orange, pink, and gray clusters are related to the place category. The implications of this research imply that improving occupational safety and health in the work environment is very important. This includes the importance of regularly maintaining and repairing public facilities to reduce the risk of accidents and damage. The research results show that the incident reporting system is able to identify potential hazards efficiently through analysis of visible patterns. In addition, these findings indicate that management needs to prioritize factors that must be addressed to improve work safety standards and contribute to reducing hazards that can impact all work potential in the organization.

5. Conclusion

The aim of this research is to analyze the pattern of reporting near accidents by company employees through the IZAT (Zero Accident Assistant Application). For this purpose, this research uses a Social Network Analysis approach to find community patterns in news and identify what keywords and nodes often appear. The results of this study show that 276 employees were actively involved in reporting findings related to near misses in the workplace, which reflects their high participation in the safety reporting program. Five categories provide a visual representation of frequently appearing words: place, response, cause, tools and condition. Network visualization shows the existence of six clusters where frequently occurring words are connected to that category. The many keywords reflect employees' sensitivity and awareness of unsafe or inappropriate conditions in their work environment. This research positively contributes to improving K3 management and mitigating potential dangers for the entire company workforce. These findings provide an in-depth evaluation of the effectiveness of the IZAT incident reporting system in detecting potential hazards in work areas based on analysis of incoming reporting patterns. In addition, it provides recommendations for focusing on hazard mitigation aspects that need to be prioritized by management to improve work safety standards in all company operational areas.

This research has several limitations, namely, the data collection period only covers four years (September 2020 - July 2023), so more is needed to analyze long-term trends related to near-miss incidents. Also, the number of near-miss incident reports entered into the IZAT system still needs to be considered more significant and uneven for all company work areas. This condition can affect the accuracy of the overall pattern analysis results. This is because the available data does not represent the population as a whole. In future research, it is recommended that the period for data collection be extended to obtain more accurate long-term trends related to near-miss incidents. In addition, the number of participants reporting near-miss incidents needs to be increased in all work areas to make the data obtained more representative.

6. References

[1] O. Felix Orikpete and D. Raphael Ejike Ewim, “Interplay of human factors and safety culture in nuclear safety for enhanced organisational and individual Performance: A comprehensive review,” *Nuclear Engineering and Design*, vol. 416, Jan. 2024, doi: 10.1016/j.nucengdes.2023.112797.

[2] K. L. Ebi *et al.*, “Extreme Weather and Climate Change: Population Health and Health System Implications,” 2021, doi: 10.1146/annurev-publhealth.

[3] C. Chaiyapet, W. Kedthongma, and W. Phakdeekul, “Risk Factors of Severity of Road Accident Injury Incidence At Kut Bak district Sakon Nakhon Province, Thailand.” [Online]. Available: https://www.researchgate.net/publication/366999098

[4] M. Niciejewska and O. Kiriliuk, “Occupational health and safety management in ‘small size’ enterprises, with particular emphasis on hazards identification,” *Production Engineering Archives*, vol. 26, no. 4, pp. 195–201, Dec. 2020, doi: 10.30657/pea.2020.26.34.

[5] N. Bugalia, Y. Maemura, and K. Ozawa, “A system dynamics model for near-miss reporting in complex systems,” *Saf Sci*, vol. 142, Oct. 2021, doi: 10.1016/j.ssci.2021.105368.

[6] S. Ki Lee and J. Park, “국내 제조업의 아차사고 발굴활동 현황 및 인식에 관한 연구 A Study on the Status and the Perception of Near Miss Reporting Activities in Domestic Manufacturing Industry,” *Journal of The Korea Convergence Society*, vol. 12, no. 12, pp. 287–294, 2021, doi: 10.15207/JKCS.2021.12.12.287.

[7] Y. Yang and H. Liu, “The effect of patient safety culture on nurses’ near-miss reporting intention: the moderating role of perceived severity of near misses,” *Journal of Research in Nursing*, vol. 26, no. 1–2, pp. 6–16, Mar. 2021, doi: 10.1177/1744987120979344.

[8] A. J. Al-Bayati, “Impact of construction safety culture and construction safety climate on safety behavior and safety motivation,” *Safety*, vol. 7, no. 2, Jun. 2021, doi: 10.3390/SAFETY7020041.

[9] M. Mambwe, E. M. Mwanaumo, W. D. Thwala, and C. O. Aigbavboa, “Evaluating occupational health and safety management strategy success factors for small-scale contractors in zambia,” *Sustainability (Switzerland)*, vol. 13, no. 9, May 2021, doi: 10.3390/su13094696.

[10] E. J. Tetzlaff, K. A. Goggins, A. L. Pegoraro, S. C. Dorman, V. Pakalnis, and T. R. Eger, “Safety Culture: A Retrospective Analysis of Occupational Health and Safety Mining Reports,” *Saf Health Work*, vol. 12, no. 2, pp. 201–208, Jun. 2021, doi: 10.1016/j.shaw.2020.12.001.

[11] M. Namian, D. Hollar, F. Taherpour, and E. Ghiasvand, “Deciphering Why ‘Cursed Construction Workers’ are More Vulnerable to Occupational Accidents,” EasyChair, Sep. 2020, pp. 491–482. doi: 10.29007/wmdk.

[12] N. Jadidi and A. A. Bazdar, “Safety Risk Management in Production Process: A case study in the automotive supplier industry,” *International Journal of Reliability, Risk and Safety: Theory and Application*, vol. 3, no. 1, pp. 85–95, Jan. 2020, doi: 10.30699/ijrrs.3.1.10.

[13] P. Pascual-Ferrá, N. Alperstein, and D. J. Barnett, “Social Network Analysis of COVID-19 Public Discourse on Twitter: Implications for Risk Communication,” *Disaster Med Public Health Prep*, vol. 16, no. 2, pp. 561–569, Apr. 2022, doi: 10.1017/dmp.2020.347.

[14] L. Zhang, X. Guo, Z. Lei, and M. K. Lim, “Social network analysis of sustainable human resource management from the employee training’s perspective,” *Sustainability (Switzerland)*, vol. 11, no. 2, Jan. 2019, doi: 10.3390/su11020380.

[15] T. Chang, S. Chi, and S.-B. Im, “Understanding User Experience and Satisfaction with Urban Infrastructure through Text Mining of Civil Complaint Data,” *J Constr Eng Manag*, vol. 148, no. 8, Aug. 2022, doi: 10.1061/(asce)co.1943-7862.0002308.

[16] C. Arifianto and V. Veritia, “SOCIAL NETWORK ANALYSIS: A COMPETITION IN INDONESIA’S FASTEST GROWING FINTECH,” *Jurnal Manajemen dan Kewirausahaan*, vol. 24, no. 1, pp. 73–80, Mar. 2022, doi: 10.9744/jmk.24.1.73-80.

[17] A. Wajahat *et al.*, “Interactively Visualize and Analyze Social Network Gephi,” in *2020 3rd International Conference on Computing, Mathematics and Engineering Technologies: Idea to Innovation for Building the Knowledge Economy, iCoMET 2020*, Institute of Electrical and Electronics Engineers Inc., Jan. 2020. doi: 10.1109/iCoMET48670.2020.9073812.

[18] G. G. Brinkmann, K. F. D. Rietveld, F. J. Verbeek, and F. W. Takes, “Real-time interactive visualization of large networks on a tiled display system,” *Displays*, vol. 73, Jul. 2022, doi: 10.1016/j.displa.2022.102164.

[19] R. G. Bhati, “A SURVEY ON SENTIMENT ANALYSIS ALGORITHMS AND DATASETS,” *Review of Computer Engineering Research*, vol. 6, no. 2, pp. 84–91, Sep. 2019, doi: 10.18488/journal.76.2019.62.84.91.

[20] M. Freire, F. Antunes, and J. P. Costa, “Applying Social Network Analysis and Data Mining Techniques to Support Decision-Making: A Case Study,” in *Frontiers in Artificial Intelligence and Applications*, IOS Press BV, Nov. 2022, pp. 41–48. doi: 10.3233/FAIA220422.

[21] G. Yunanda, D. Nurjanah, and S. Meliana, “Recommendation System from Microsoft News Data using TF-IDF and Cosine Similarity Methods,” *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 1, Jun. 2022, doi: 10.47065/bits.v4i1.1670.

[22] S. Sakthi Vel, “Pre-Processing techniques of Text Mining using Computational Linguistics and Python Libraries,” in *Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021*, Institute of Electrical and Electronics Engineers Inc., Mar. 2021, pp. 879–884. doi: 10.1109/ICAIS50930.2021.9395924.

[23] A. A. Ojugo and A. O. Eboka, “Memetic algorithm for short messaging service spam filter using text normalization and semantic approach,” *International Journal of Informatics and Communication Technology (IJ-ICT)*, vol. 9, no. 1, p. 9, Apr. 2020, doi: 10.11591/ijict.v9i1.pp9-18.
[24] B. Göschlberger and D. Deliu, “BiRank vs PageRank: Using SNA on Company Register Data for Fiscal Risk Prediction,” Sep. 2022, doi: 10.1109/SNAMS53716.2021.9732111.
[25] H. Pang, J. Liu, and J. Lu, “Tackling fake news in socially mediated public spheres: A comparison of Weibo and WeChat,” *Technol Soc*, vol. 70, Aug. 2022, doi: 10.1016/j.techsoc.2022.102004.
[26] J. Butar Butar, S. Lusa, S. Handayani, and A. A. Yusuf, “Analyzing Abstention Discourse in Presidential Elections: Knowledge Discovery in X Using ML, LDA and SNA,” vol. 14, no. 3, 2024.
[27] U. A. Marovac, A. M. A. Avdi´c, and A. B. Ljaji´cljaji´c, “Creating a Stop Word Dictionary in Serbian,” 2021.
[28] F. Aghware *et al.*, “Sentiment Analysis in Detecting Sophistication and Degradation Cues in Malicious Web Contents,” vol. 38, 2023, [Online]. Available: https://www.researchgate.net/publication/374875030
[29] P. K. Hota, B. Subramanian, and G. Narayanamurthy, “Mapping the Intellectual Structure of Social Entrepreneurship Research: A Citation/Co-citation Analysis,” *Journal of Business Ethics*, vol. 166, no. 1, pp. 89–114, Sep. 2020, doi: 10.1007/s10551-019-04129-4.
[30] A. Ullah, B. Wang, J. F. Sheng, J. Long, N. Khan, and Z. J. Sun, “Identification of nodes influence based on global structure model in complex networks,” *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-84684-x.
[31] S. Azad and S. Devi, “Tracking the spread of COVID-19 in India via social networks in the early phase of the pandemic,” *J Travel Med*, vol. 27, no. 8, pp. 1–9, 2021, doi: 10.1093/JTM/TAAA130.
[32] R. F. Betzel, J. Faskowitz, and O. Sporns, “Living on the edge: network neuroscience beyond nodes,” Nov. 01, 2023, *Elsevier Ltd*. doi: 10.1016/j.tics.2023.08.009.
[33] N. A. Prabowo, B. Pujiarto, F. S. Wijaya, L. Gita, and D. Alfandy, “Social Network Analysis for User Interaction Analysis on Social Media Regarding E-Commerce Business,” *International Journal of Informatics and Information System*, vol. 4, no. 2, pp. 95–102, 2021.
[34] F. Li, W. Liu, and K. Bi, “Exploring and visualizing spatial-temporal evolution of patent collaboration networks: A case of China’s intelligent manufacturing equipment industry,” *Technol Soc*, vol. 64, Feb. 2021, doi: 10.1016/j.techsoc.2020.101483.
[35] A. S. Ferreira, M. Sacomano Neto, S. E. A. Candido, and G. M. Ferratti, “Network Centrality and Performance: Effects in the Automotive Industry,” *Revista Brasileira de Gestao de Negocios*, vol. 23, no. 4, pp. 677–695, 2021, doi: 10.7819/rbgn.v23i4.4132.
[36] R. Schiff, K. Arnold, and A. Wilkinson, “Using social network analysis to understand the impact of systems integration efforts: a case study from Thunder Bay,” *Cogent Soc Sci*, vol. 10, no. 1, 2024, doi: 10.1080/23311886.2024.2320463.
[37] A. Alamsyah, D. P. Ramadhani, and L. S. Mulyani, “Rise or fall? Discovering the global world trade network rise and fall under major situations,” *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 9, no. 1, Mar. 2023, doi: 10.1016/j.joitmc.2023.100009.