EFFICIENT SAFETY INFORMATION RETRIEVAL ON CONSTRUCTION SITES: A PRELIMINARY METHODOLOGY

Hao Zhang[1], Seokho Chi[2*], Jay Yang[3] and Ting-Kwei Wang[4]

[1]School of Civil Engineering and Built Environment, Queensland University of Technology, Brisbane, QLD 4000, Australia; email: zhang.hao@student.qut.edu.au
[2*]**Corresponding author**, School of Civil Engineering and Built Environment, Queensland University of Technology, Brisbane, QLD 4001, Australia; PH (+61) 7 3138 9988; FAX (+61) 3138 1170; email: seokho.chi@qut.edu.au
[3]School of Civil Engineering and Built Environment, Queensland University of Technology, Brisbane, QLD 4000, Australia; email: j.yang@qut.edu.au
[4]Faculty of Construction Management and Real Estate, Chongqing University, Chongqing, 400045, China; email: tingkwei@cqu.edu.cn

ABSTRACT

IT-supported field data management benefits on-site construction management by improving accessibility to the information and promoting efficient communication between project team members. However, most of on-site safety inspections still heavily rely on subjective judgment and manual reporting processes and thus observers' experiences often determine the quality of risk identification and control. This study aims to develop a methodology to efficiently retrieve safety-related information so that the safety inspectors can easily access to the relevant site safety information for safer decision making. The proposed methodology consists of three stages: (1) development of a comprehensive safety database which contains information of risk factors, accident types, impact of accidents and safety regulations; (2) identification of relationships among different risk factors based on statistical analysis methods; and (3) user-specified information retrieval using data mining techniques for safety management. This paper presents an overall methodology and preliminary results of the first stage research conducted with 101 accident investigation reports.

1. Introduction

Accurate and prompt field information can support efficient and effective management of construction projects (Taneja *et al.*, 2011). In recent years, IT-based solutions such as mobile computing and smart sensing technologies have improved accessibility to the project information and enhanced communication between project team members while maximizing project time and cost savings (Peansupap and Walker, 2006). Saidi et al. (2002) applied handheld computers for six construction field activities including punch-listing, materials tracking, access to a material and safety database, requests for information (RFI), drawing access and quantity surveying. The case study results demonstrated the potential for time saving and quality improvement of construction operations. Kim et al. (2011) presented a location-based construction site management system using a mobile computing communication platform. The system architecture comprised a site management module that explained the location information of both construction activities and the resources (i.e. materials and workers) allocated to the activities, and a construction drawing sharing module that provided an easy access to construction drawing. Torrent and Caldas (2009) employed Radio Frequency Identification (RFID) technology and enabled automated material identification, localization and tracking while reducing human errors, improving material management efficiency and thus resulting in productivity benefits. Akinci *et al.* (2006) and Taneja *et al.* (2011) successfully collected as-built building information using laser scanning technologies, compared them with the as-planned information, and eventually contributed to efficient project progress monitoring achieving cost and quality benefits.

Construction safety management is also information driven and safety inspection is one of the key safety management elements that assesses working conditions on a construction site, identifies existing risks and suggests actions for mitigation. However, in the conventional safety management practices, the on-site inspection still heavily relies on subjective judgment and manual reporting processes and thus the observer's experience and competency often determine the inspection performance (Fung *et al.*, 2010). A less-skilled safety inspector may have difficulties in identifying and controlling on-site safety risks and even a highly-experienced inspector can miss some of relevant risk information. Therefore, IT-based information communication currently applied to the other aspects of construction management practices previously discussed will also benefit safety management by providing safety managers with an easy access to supportive safety information.

In this regard, some researchers investigated automated data collection and analysis approaches for real-time safety assessment and injury prevention on a construction site. Teizer *et al.* (2010) proposed an autonomous system that tracked the movement of workers and heavy equipment and alerted them when they were in a possible collision status. Chi and Caldas (2011) presented an exploratory method for automated object identification using standard video cameras on construction sites. It showed future application potential of automated monitoring of construction safety. Chi and Caldas (2012) developed an automated warning mechanism against dangerous earthmoving conditions. In this study, construction resources were automatically tracked using 3D video cameras and the research output promoted real-time safety assessment on construction sites. Despite their great efforts, they are still in a preliminary testing stage away from the practical implementation due to the complicated and unpredictable nature of construction sites. Other researchers proposed safety management databases, as more practical utilization tools, that incorporated historical accident information to provide reference data for risk identification and control (Carter and Smith, 2006; Goh and Chua, 2009). However, they were generally a simple representation of safety information without strategic information retrieval concepts not only explaining which information is important to be considered and how to extract relevant information for safety assessment but also providing a user with specific step-by-step inspection guidelines.

Thus, the research presented in this paper aims to investigate an efficient information retrieval methodology to support strategic safety inspection. The proposed methodology targets (1) building a comprehensive safety database which contains information of risk factors, accident types, impact of accidents and safety regulations; (2) identifying relationships among different risk factors based on statistical analysis methods; and (3) enabling user-specified information retrieval using data mining techniques for safety management. This methodology will be then used to develop an intelligent electronic inspection checklist for the safety inspectors and provide them with easily accessible information and step-by-step inspection guidelines to improve the performance of risk identification and control. The safety risks in this research are limited to physical, on-site operational conditions in building construction. It does not include safety culture and personal cognitive behaviors.

2. Research Methodology

To achieve the project aim, three research objectives will need to be explored, studied and answered. *Stage 1*: To begin with, safety accident investigation reports and safety regulations will be thoroughly reviewed and analyzed to identify risk factors. A comprehensive safety accident database will be built with the identified risk factors and related safety regulations and best practices. Although all the information in the finalized accident database will be grouped into construction operation by operation (e.g. roofing, earthmoving or concrete works), risk factors of each individual accident type (e.g. falling from the height or struck by heavy equipment) will be separately analyzed first since most of accident investigation reports can be easily sorted by different accident types. They will be then combined together and reorganized to represent site conditions of each operation. *Stage 2*: The next step is to design the methodology to retrieve information efficiently from the database. When a user inputs some site conditions into the information retrieval system using the provided inspection checklist, relevant safety information and following inspection items should be automatically instructed step by step based on the methodology. As this study proposes that the relationship between different risk factors can contribute to efficient information retrieval, a safety risk relationship diagram will be developed by using statistical analysis methods and verified through the interview with industry experts. *Stage 3*: Based on the developed relationship diagram, an information retrieval methodology will be further developed and the structured computer programming will implement the methodology. Mobile computing technology will be utilized for the methodology implementation. The methodology will be then tested and validated using historical accident data and through case studies. The research findings will finally derive conclusions and recommendations. This paper presents preliminary results of the first two stages including accident investigation report review, risk factor identification and relationship analysis through statistical analysis methods.

3. Preliminary Risk and Relationship Analysis

3.1 Description of the data

The authors first thoroughly reviewed 101 accident investigation reports issued by the U.S. Occupational Safety and Health Administration (OSHA), the U.S. National Institute for Occupational Safety and Health (NIOSH) and the Workers' Compensation Board of British Columbia in Canada (WorkSafeBC). All of them referred to the falling from the height accidents that occurred during the building construction projects. The authors then identified 25 risk factors associated with the falling accidents by comparing the reported data with the literature review findings and grouped them into five categories: (1) work being done, (2) victim information, (3) working environment, (4) safety risk mitigation and (5) time information. The work being done category represented task-related information including type of work, hazard sources, working height, resource information and number of workers. The victim information category included age, eye vision and worker's movement factors. The working environment explained surrounding working conditions including weather, noise level, ground stability and surface condition, and others. The safety risk mitigation category included safety protection elements such as guardrail systems, safety nets and personal harness systems. Lastly, the time category explained date and seasonal information of the accident occurred. Table 1 shows examples of three risk categories and their individual observation frequency counted from the investigation reports.

Table 1. Examples of risk category and the frequency of risk observations

Category	Factors	Observations	Frequency
Work being done	Type of work	Formwork and reinforcing steel	2%
		Structural concrete construction	1%
		Floor construction	5%
		Roof work	37.60%
		Masonry	4%
		Carpentry work	6.90%
		Painting	5%
		Electrical work	5%
		Equipment installation	3%
		Iron/steel erection	11.90%
		Scaffolding	3%
		Exterior work	5.90%
		Demolition	2%
		Cleaning	2%
		Other work or unknown	5.90%
	Hazard source	Leading edges	6.90%
		Holes	11.90%
		Formwork and reinforcing steel	1%
		Roof work	32.70%
		Wall openings	1%
		Perimeter scaffold	13.90%
		Suspended scaffold	5%
		Elevating work platforms	10.90%
		Ladder	5.90%
		Fragile surface	1%
		Structure steel	9.90%
	Height	6-26ft	56.40%
		27-47ft	21.80%
		48-68ft	11.90%
		69-89ft	3.00%
		>=90ft	5.00%
	Equipment	No equipment involved	39.60%
		Machine	9.90%
		Handheld devices	50.50%
	Material quality	Normal	93.10%
		Unqualified	6.90%
	Number of workers	Only the victim	31.70%
		2--5	64.40%
		>5	4.00%
Victim information	Age	<18	5.90%
		18-24	16.80%
		25-34	25.70%
		35-44	22.80%
		45-54	12.90%
		>=55	11.90%
	Vision	Normal	97.00%
		Impacted	3.00%
	Worker status	Whole body moving, not backwards	53.50%
		Whole body moving, backwards	12.90%
		Partial body moving or stationary	33.70%
	Work with handling	Yes	47.50%
		No	52.50%
Working Environment	Weather	Normal	73.30%
		Windy	5.90%
		Snowy	2.00%
		Rainy	1.00%
		Cold	5.00%
		Cloudy	2.00%
		Sunny	2.00%
		Indoor	8.90%
	Surrounding setting	Normal	94.10%
		Crowded	5.00%
		Noisy	1.00%
	Ground stability	Normal	95.00%
		Unfavorable	5.00%
	Surface condition	Normal	61.40%
		Sloping	8.90%
		Slippery	5.90%
		Unstable	9.90%
		Fragile	9.90%
		Uneven	4.00%
	Regular inspection	Yes	32.70%
		No	31.70%
	Onsite monitoring	Yes	27.70%
		No	33.70%

* Some "unknown" observations were excluded for the analysis.

3.2 Relationship analysis

The correlations between every two risk factors were then statistically analyzed. Fisher's exact test (Fisher, 1954; Agresti, 1992) was implemented for the analysis using the SPSS software because this test is often used as an alternative to the Pearson Chi-Square when one or more cells contain a small number of observations. The null hypothesis of Fisher's exact test is that two variables to be examined are independent with each other. Let us suppose two variables *X* and *Y* have *m* and *n* observations respectively. An $m \times n$ matrix can be formed, in which the entries a_{ij} represent the frequency of observations. *i* increases by *m* and *j* increases by *n*. For instance, if the frequency of the entry at the first row and the first column is 12, $a_{11} = 12$. R_i and C_j represent the row and column sums, and *N* is the total sum of R_i or the total sum of C_j. The sum of R_i equals the sum of C_j. The Fisher's exact test then calculates the condition probability of the matrix with this information and defines it as P_{cutoff} (eq. 1):

$$P_{cutoff} = \frac{(R_1!\,R_2! \ldots R_m!)(C_1!\,C_2! \ldots C_n!)}{N!\,\prod_{ij} a_{ij}!} \quad (eq.\ 1)$$

The test then calculates the conditional probability of every possible matrix with the fixed R_i and C_j values based on the same equation. This individual *p*-value is compared with P_{cutoff} and the sum of *p*-values less than or equal to P_{cutoff} becomes the representative *p*-value of the test. If this *p*-value becomes less than 0.05, the null hypothesis should be rejected, which means that there is a significant correlation between the two variables. In this study, Fisher's exact test identified the total 36 pairs of interrelated factors (Table 2).

Table 2. Correlated factors identified by Fisher's exact test

Factor1	**Factor2**	**Fisher's exact test p-value**
Type of work	Hazard source	0
	Worker status	0.042
	Regular inspection	0.023
	Safety net system	0.006
	Month of accident	0
Hazard source	Equipment	0
	Worker status	0.008
	Safety net system	0.004
	Month of accident	0
Height	Guardrail system	0.005
	Personal fall arrest system	0.006
	Other fall protection system	0.006
Equipment	Material quality	0.002
	Worker status	0.038
	Work with handling	0.002
	Onsite monitoring	0.028
	Safety net system	0.002
Material quality	Onsite monitoring	0.027
	Safety net system	0.042
Number of workers	Personal fall arrest system	0.001
Age	Vision	0.033
Vision	Weather	0.027
Worker status	Ground stability	0.006
	Cover	0.012
	Month	0.02
Work with handling	Weather	0.01
Surrounding setting	Safety net system	0.028
Ground stability	Guardrail system	0.02
Surface condition	Guardrail system	0.018
	Cover	0.019
Regular inspection	Onsite monitoring	0
	Month of accident	0.005
Guardrail system	Cover	0
	Safety net system	0.005
Safety net system	Personal fall arrest system	0
Personal fall arrest system	Other fall protection system	0.018

* A p-value less than or equal to 0.05 indicates a significant relationship.

The authors qualitatively analyzed these test findings by comparing them with the original accident investigation reports and here are examples of the identified relationships:

- *"Type of work" is correlated to "Hazard source"*. e.g. the holes on the floor (hazard source) resulted in 60% of falling accidents during the floor construction (type of work) and scaffolding problems and the improper elevation control of work platforms (hazard source) accounted for 80% of accident during the painting (type of work).
- *"Type of work" is correlated to "Safety net systems"*. e.g. the safety nets sometimes become more important depending on the type of works. The roofing and scaffolding works usually require more safety nets than the floor and cleaning works.
- *"Falling protection systems" are correlated to "Height"*. e.g. the more working elevation increases, the more safety managers pay attend to protective equipment such as guardrail systems or personal fall arrest systems. Additional falling protection systems need to be installed on top of the conventional systems for the working elevation higher than 90ft.
- *"Material quality" is correlated to "On-site monitoring"*. e.g. the accident reports showed 71.4% of poor quality materials were related to lack of proper on-site monitoring practices.
- *"Working with material handling" is correlated to "Weather"*. e.g. the working with material handling means that the worker was handling materials when the accident occurred. 10.4% of this type of accident occurred during the windy weather since the strong wind made workers lost their balance during their work.

The authors then drew a risk relationship diagram to illustrate the analysis results by connecting two correlated risks and investigated a possible chain of risk events for accident generation (Figure 1). For instance, the poor material quality can be inspected through the on-site monitoring and the performance of the on-site monitoring is highly related to the regular inspection practices. Additionally, different regular inspection approaches are determined by different work types. As another example, the surface condition determines the applicability of the surface cover and the cover limits the movement of workers. The worker's movement is also related to different work types and different kinds of equipment used for the operations.

4. Conclusion

This paper presented the methodology for efficient safety information retrieval. The preliminary analysis that was conducted with 101 accident investigation reports identified 25 risk factors inherent in falling from the height accidents, determined 36 inter-related risk pairs through Fisher's exact test and developed a risk relationship diagram using the research findings. This information is expected to guide safety inspectors to check all the relevant risks during the inspection process when one risk is identified. In this way, higher efficiency and effectiveness of safety inspection might be achieved.

The research team is currently conducting similar analysis for different accident types. Further verification should be performed by seeking industry feedback on the identified risk factors and their relationships to make the diagrams more practical sense. A comprehensive operation-based safety database will be then constructed with the holistic relationship diagram by consolidating all risk factors and their relationships identified. They will be integrated with data mining techniques and eventually develop an electronic safety inspection checklist.

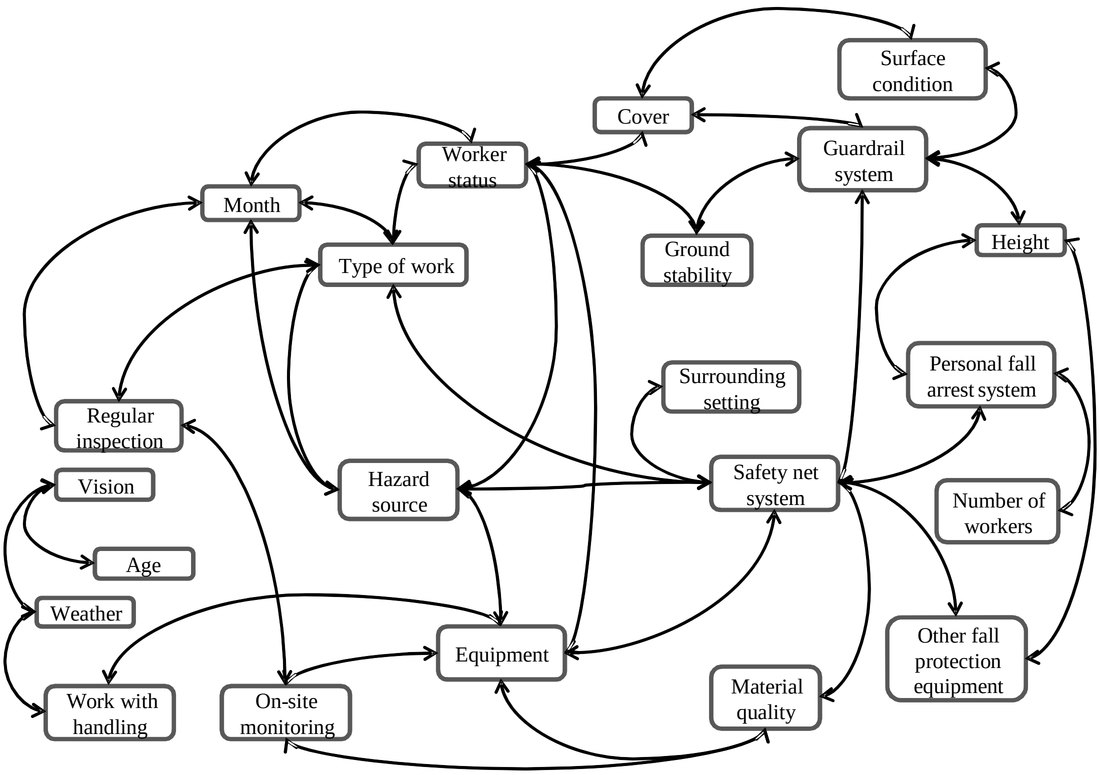

Figure 1. Risk relationship diagram of the falling accident

References

Agresti, A. (1992). “A survey of exact inference for contingency tables.” *Statistical Science*, 7(1), 131-153.

Akinci, B., Boukamp, F., Gordon, C., Huber, D., Lyons, C., and Park, K. (2006). "A formalism for utilization of sensor systems and integrated project models for active construction quality control." *Automation in Construction*, 15(2), 124-138.

Carter, G., and Smith, S. D. (2006). "Safety hazard identification on construction projects." *Journal of Construction Engineering and Management*, 132(2), 197-205.

Chi, S., and Caldas, C. H. (2012). "Image-based safety assessment: automated spatial safety risk identification of earthmoving and surface mining activities." *Journal of Construction Engineering and Management*, 138(3), 341-351.

Chi, S. and Caldas, C. H. (2011). "Automated object identification using optical video cameras on construction sites." Computer-Aided Civil and Infrastructure Engineering, Special Issue on Advances in Construction Automation, 26(5), 368-380.

Fisher, R. A. (1954). *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh, UK.

Fung, I. W. H., Tam, V. W. Y., Lo, T. Y., and Lu, L. L. H. (2010). "Developing a Risk Assessment Model for construction safety." *International Journal of Project Management*, 28(6), 593-600.

Goh, Y. M., and Chua, D. K. H. (2009). "Case-based reasoning for construction hazard identification: Case representation and retrieval." *Journal of Construction Engineering and Management*, 135(11), 1181-1189.

Kim, C. Lim, H., and Kim, H. (2011). "Mobile computing platform for construction site management." *Proceedings on the 2011 International Symposium on Automation and Robotics in Construction (ISARC)*, Seoul, Korea.

Mehta, C. R., and Patel, N. R. (2010). "IBM SPSS Exact Tests." 18-23.

Peansupap, V., and Walker, D. H. T. (2006). "Innovation diffusion at the implementation stage of a construction project: A case study of information communication technology." *Construction Management and Economics*, 24(3), 321-332.

Saidi, K. S., Haas, C. T., and Balli, N. A. (2002). "The value of handheld computers in construction." *Proceedings on the 2002 International Symposium on Automation and Robotics in Construction*, National Institute of Standards and Technology, Gaithersburg, Maryland.

Taneja, S., Akinci, B., Garrett, J. H., Soibelman, L., Ergen, E., Pradhan, A., Tang, P., Berges, M., Atasoy, G., Liu, X., Shahandashti, S. M., and Anil, E. B. (2011). "Sensing and field data capture for construction and facility operations." *Journal of Construction Engineering and Management*, 137(10), 870-881.

Teizer, J., Allread, B. S., Fullerton, C. E., and Hinze, J. (2010). "Autonomous pro-active real-time construction worker and equipment operator proximity safety alert system." *Automation in Construction*, 19(5), 630-640.

Torrent, D. G., and Caldas, C. H. (2009). "Methodology for Automating the Identification and Localization of Construction Components on Industrial Projects." *Journal of Computing in Civil Engineering*, 23(1), 3-13.