Safety Science 70 (2014) 89–98

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/ssci

Identifying patterns of safety related incidents in a steel plant using association rule mining of incident investigation reports

Abhishek Verma [a], Sudha Das Khan [a], J. Maiti [a,*], O.B. Krishna [b]

[a] Department of Industrial & Systems Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India
[b] Corporate Safety, Safety, Health and Environment, Tata Steel, Jamshedpur, India

ARTICLE INFO

Article history:
Received 10 November 2013
Received in revised form 13 April 2014
Accepted 13 May 2014

Keywords:
Incident investigation
Association rules
Incident patterns
Steel plant

ABSTRACT

The aim of this paper is to find out the patterns of incidents in a steel plant in India. Occupational incidents occur in steel plant mainly in form of injury, near miss, and property damage or in combination. Different factors are responsible for such incidents to occur. An incident investigation scheme is proposed. Association rule mining approach is used to discover cause-and-affect patterns (rules) using 843 incidents. Thirty-five meaningful association rules are extracted using three criteria, support (S), confidence (C) and lift (L). For example, the results show that unsafe acts done by others are more frequent in injury cases ($S = 4.86\%$, $C = 78.8\%$, $L = 2.3$). Similarly, one of the SOP (standard operating procedures) related rule: 'SOP required, available, adequate but not complied' led to property damage ($S = 11.03\%$, $C = 49.2\%$, $L = 1.525$). Another useful rule 'SOP required, available but inadequate, followed' led to near miss ($S = 1.66\%$, $C = 38.89\%$, $L = 1.163$). It is also found that for slip, trip and fall incidents, workers working alone ($S = 3.91\%$, $C = 76.74\%$, $L = 2.239$) or in a group ($S = 3.20\%$, $C = 75.00\%$, $L = 2.188$) does not make much difference. The findings pinpoint the areas of improvement such as inadequate SOPs, non-compliance of SOPs, training, and slip, trip and fall prevention to minimize incidents.

1. Introduction

Steel manufacturing is one of the most hazardous industries because of its complex socio-technical system. In one hand, it uses high technology and on the other hand, it is labor intensive. It has all components of process safety and personal safety. Managing safety in steel industry is a daunting task. According to domino theory proposed by Heinrich (1959), maximum industrial incidents occur due to unsafe acts, unsafe conditions, or both; which are controllable. Reason's Swiss Cheese model depicts that an hazard becomes an accident, when a series of events in between align together thus creating a path from hazard to accident (Reason, 1990). Identification and quantification of such paths are utmost important for preventing accidents to occur.

The analysis of accident data is as old as the development of Poisson distribution. Fatal accident being a rare event fits Poisson distribution very well. Over the years, several statistical models have been used to explore factors causing incidents. Khanzode et al. (2012) have done a review work on accident causation theories generation wise; like 1st generation theories on accident proneness (Greenwood and Woods, 1919), 2nd generation theory about Domino theories (Heinrich, 1959), 3rd generation injury epidemiology theories (Haddon et al., 1964) and 4th generation theories on system approach (Trist and Bamforth, 1951). Some pioneering works on accident data analysis includes (Cooper, 2000; Shankar and Mannering, 1995; Maher and Summersgill, 1996).

In recent years, the analysis of incident/accident data using data mining techniques and algorithms has been gaining much attention among researchers (Pande and Abdel-Aty, 2009; Liao and Perng, 2008; Cheng et al., 2013; Arunraj et al., 2013). Recently, Maiti et al. (2014) proposed a methodology to determine safety rules for derailments in a steel plant using corresponding analysis. Most of industrial incidents occur due to lack of knowledge, lack of standard operating procedure and its compliance, insufficient training, etc. Researchers are trying to discover the causes of incidents which can be used to improve safety management system in industry. Many data mining techniques such as support vector machines, classification and regression trees and Bayesian networks have been used to identify hidden patterns and structures in a large amount of data consisting of various factors associated with incidents or accidents.

The association rule mining (Agrawal et al., 1993) has been used to analyze incident data to get rules for incident patterns. For example, in road incident, crash data analysis was done to find

* Corresponding author. Tel.: +91 3222283750; fax: +91 3222255303.
E-mail addresses: abhishekverma.cs@gmail.com (A. Verma), essd.iit@gmail.com (S.D. Khan), jhareswar.maiti@gmail.com (J. Maiti), obkrishna@gmail.com (O.B. Krishna).

http://dx.doi.org/10.1016/j.ssci.2014.05.007
0925-7535/

out the contributory factors and their interdependency (Pande and Abdel-Aty, 2009; Montella, 2011; Montella et al., 2012); in railway, accidents data analysis was done to discover the factors and relationship between them (Mirabadi and Sharifian, 2010); and in construction industry, incident data was analyzed to find out cause and effect relationships of occupational incidents (Liao and Perng, 2008; Cheng et al., 2010). One of the well-received algorithm for association rule mining is apriori algorithm (Agrawal and Ramakrishnan, 1994). Apart from apriori algorithm there are different algorithms, techniques and approaches that have been used for mining association rules are, FP-Growth algorithm (Han et al., 2000), Eclat algorithm (Zaki, 2000), Maxclique algorithm (Zaki et al., 1997), partitioning technique (Savasere et al., 1995), sampling approach (Toivonen, 1996), continuous association rule mining algorithm (CARMA) (Hidber, 1999), vertical itemset partitioning for efficient rule extraction algorithm (VIPER) (Shenoy et al., 2000), aprioriTid and aprioriHybrid algorithms (Agrawal and Ramakrishnan, 1994). Although there are several studies on association rule mining conducted, but in safety data analysis, literature is scanty.

Most of the leading organizations across the world have their own incident or accident investigation and reporting system. The primary task of incident investigation is to identify "what", "where", "when", "why" and "how" incidents have happened. It also aims to find out the root causes behind a particular incident to take corrective actions to avoid its reoccurring. The key information regarding incidents is recorded in investigation reporting system. Appropriate factors are used in incident investigation reporting according to industry type (Brazier, 1994). Incident investigation reporting has been used by various industries to find the causal factors and near misses to improve safety performance (Ji and Zhang, 2012; Okstad et al., 2012; Nesmith et al., 2013; Jones et al., 1999). Basso et al. (2004) proposed an incident investigation database system to record factors causing incidents as well as corrective actions that allow monitoring safety performance. Oktem et al. (2010) also proposed a model to design near miss management system by a framework from event identification to solution implementation. Gnoni et al. (2013) proposed similar type of model taking benefit of lean thinking. While the incident reporting system is somehow well managed but it is still flawed from two counts: (i) lack of process approach incorporating workflow information among key stake holders like supervisor, department head, safety professional, etc. and (ii) data gathered is hardly analyzed in the way it is purposed. In this study, we attempt to provide a scheme for incident investigation followed by incident data analysis.

This study starts with describing the incident data and explaining the proposed incident investigation process (Section 2.1). It comprises the most prominent combination of factors leading to one of three types of incidents: injury, near miss, and property damage. Association rule mining using apriori algorithm is employed to extract incident patterns in Section 2.2. Data codification and analysis is given in Section 2.3. The results obtained and its practical implications are given in Section 2.4. Inference of the study to take managerial decisions has been discussed in Section 3. Finally, conclusions of the study with future scope of research are given in Section 4.

2. Methods

2.1. Incident investigation process

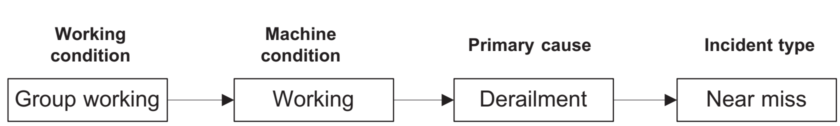
The study was conducted in coke, sinter and iron (CSI) division of a steel manufacturing company in India. To ensure safety, the company uses online safety management system (SMS) which is available on local network inside the organization. This system provides worksite observations, fatality risk control programs, incident investigations, and job cycle checks. For this study, incident data recorded in the SMS is considered. The SMS generates excel report of the incidents logged on. In our study we have proposed the workflow for incident investigation as shown in Fig. 1. All incidents are investigated, and incident investigation reports are created. On the basis of investigation, actions are taken to improve safety (both unsafe acts and/or conditions). Any employee involved or witnessing an incident can report the same to the corresponding supervisor of the department. Supervisor then log on the incident investigation module of the SMS and fills the information fields. Then the severity of the incident is assessed considering all the hazardous elements and conditions prevailed during the incident. Depending upon the severity, risk score is given to that particular incident. If risk score is higher than a threshold limit, a prespecified value determined by the organization, it is considered as high priority incident and is then sent to the head of department (HOD) for further consideration. Low priority incident cases are taken care at the supervisory level. To handle high priority incident cases, HOD forms an investigation team of specialists to further investigate the incident scenario and explore the causal factors. After the investigation, recommendations are released by the team for implementation. Whether recommendations are correctly implemented or not is verified by safety professional. Presently, 15 information fields are generated. For our study ten out of fifteen factors (fields) have been considered as per the discussion with safety expert for in-depth analysis. Proper description for those ten factors with related information is given in Table 1.

2.2. Association rule mining

2.2.1. Illustrative examples

Example 1: A group of workers tries to couple three loaded boxes with a loco, which already is attached with four empty boxes. One of the loaded boxes is derailed because of hard pushing. No one is injured, i.e., near miss happens. But this might be lead to a major accident. While analyzing the incident, the following rules may be generated:

Rule 1:



Rule 2: