Journal of Loss Prevention in the Process Industries 56 (2018) 104–118

Contents lists available at ScienceDirect

Journal of Loss Prevention in the Process Industries

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

A new method to study the performance of safety alarm system in process operations

Hangzhou Wang[a,b], Faisal Khan[a,*], Majeed Abimbola[a]

[a] Centre for Risk, Integrity and Safety Engineering (C-RISE), Faculty of Engineering and Applied Science, Memorial University, St. John's, NL, A1B 3X5, Canada
[b] China Petroleum Planning and Engineering Institute (CPPEI), China National Petroleum Corporation, Beijing, 100083, China

ARTICLE INFO

Keywords:
Alarm data analysis
Alarm flood management
DCS data analysis
Graphical method
Weighted scoring function
Bayesian network

ABSTRACT

A new scoring function to learn the dependence structure from a process alarm dataset is proposed here. Based on the new function, a Bayesian network based method is developed to analyze alarm system performance. The proposed method is composed of four features: i) identifying essential variables from alarm data, ii) learning the structure of the Bayesian network from the alarm data; iii) learning the quantitative dependence of the variables in the Bayesian network, and iv) quantitatively describing the strength among the dependent variables. The proposed method helps to describe alarm variables relationships better, enable alarm diagnostic to identify the root causes, and study the variables dependence strength and finally improve the alarm system performance. The proposed method is first explained with a simple example, and further application is demonstrated with a case study of an industrial distributed control system-based alarm system.

1. Introduction

Distributed control systems (DCSs) are widely used in the chemical process to ensure process safety and maintain product quality. Moreover, an alarm system is introduced to help operators take proper actions (Sheridan, 2006). Chemical processes are complicated, which increases the difficulty for an operator to deal with alarm flooding (Noyes and Bransby, 2001). It is reported that alarm floods are inevitable (Zhu et al., 2014).

Alarm flooding occurs because the operator cannot deal with the alarms in a timely way (Skjerve and Bye, 2011). The operator may not be able to take appropriate corrective actions in time as he could not analyze the cause and consequence of the alarms. Either the operator was not adequately trained, or the operator was unfit for the task assigned to him. In fact, alarm flooding is the root cause of incidents such as the disasters in the Three Mile Island Nuclear plant (Stanton and Baber, 1995), the Texaco Pembroke Refinery (HSE and HEALTHSAFETY EXECUTIVE, 1997; Adnan et al., 2011), and others. In the last accident, there were 275 alarms within eleven minutes before the explosion. Although alarm flooding cannot be the only root cause of any disaster, failure of understanding, wrongful interpretation and or non-acknowledging the alarms in time and human error may also lead to failure to take corrective actions in time which can be the cause of such disasters. Moreover, Interlock enabled emergency shut down systems should be able to protect the plants under such situations. However, it was difficult for operators to identify the root causes of alarm flooding and deal with them in the emergency situation in this disaster.

To reduce or avoid alarm flooding more efforts should be taken to build an alarm management system. It would provide: i) the concise and precise representation of the alarm interface operators; ii) the relationships between the monitored variables, including dynamical characters such as the diagnosed root cause; iii) the alarm management system's requirement to learn from data automatically to cope with the variances in time. A concise and clear alarm management method will help to convey an alarm message effectively. Studying the relationships of alarms will determine the root causes of an abnormal event. This can enable a swift action to bring the system back to its normal state. Learning from data enables the alarm management system to reveal the actual state of the system and keep the state updated with time.

For the first point, the graphical and visualized interface is better for conveying messages to operators. Recent research has developed a visualized and graphical method for presenting alarms' data for alarm system analysis. Yang et al., 2011, 2012a applied the Gaussian kernel method to capture and visualize the correlation information from historical alarm data. It presented correlations between only two variables. Peng et al. (2015) used a probabilistic signed digraph (PSDG) to present the correlations of monitored variables alone.

For the second point, the knowledge-based model approaches have

* Corresponding author.
E-mail address: fikhan@mun.ca (F. Khan).

https://doi.org/10.1016/j.jlp.2018.08.017
Received 26 May 2018; Received in revised form 15 July 2018; Accepted 20 August 2018
Available online 22 August 2018

been widely used to study the causes of these alarms. Wang et al. (2016) provided an overview of industrial alarm systems and pointed out that alarm systems play critically important roles for the safe and efficient operation of modern industrial plants. Yang et al., 2012b, 2013 studied the correlation analysis for bivariate alarm signals to indicate whether two alarm signals were correlated and had a cause-effect direction, but only correlations between two variables were considered. Duan et al. (2013), Su et al. (2017) and Yu et al. (Yu and Yang, 2015) proposed a direct transfer entropy approach to detect a direct information flow pathway from one variable to another. Yang et al. (2014) built a topological model considering the connectivity and causality for root cause analysis, but the model failed to consider alarm states. Noda et al. (Noda et al., 2011) analyzed event correlation of event log data to reduce the number of alarms and operations. However, root causes were not studied. Moreover, current knowledge-based alarm systems have several drawbacks; they are language dependent, hardly scalable and their development is time and effort consuming (Abele et al., 2013).

For the third point, the chemical processes frequently vary due to the complex nonlinearity of the processes and the poor operating setpoints that are prone to fall into unstable states. These states change rapidly in a process. The ability of an alarm management system to update the model in time is required to analyze an alarm system better. Typically, there may be multiple steady states or multimode, and some operating points may be unstable, where bifurcations or oscillations have taken place, and the existence of the singularity point makes the processes more complicated. Furthermore, modern plants are a complex network of devices which are subject to frequent modifications. As a result, the knowledge-based method is not quick enough to thoroughly cover these complex characteristics, or deal with the frequent modifications in time.

Focused on the above points Bayesian networks can be employed to design an alarm management system. The Bayesian networks is helpful because i) the directed acyclic graphical interface of variables clearly represent the qualitative information in the alarm system; ii) the back diagnosis and forward inference ability can cope with the quantitative information in the alarm system; and iii) the data-driven modeling ability make the alarm system interact with variances in time. As a result, the Bayesian network can cover these aspects simultaneously. In this paper, a new method for Bayesian network learning from an alarm dataset is proposed to handle the alarm time length in a particular state more effectively. With the new approach, a new alarm management system methodology investigates these monitored variables to study the detailed dynamic characteristics to improve alarm system performance further.

The following section presents preliminary concepts of the Bayesian network; in section 3 a Bayesian network based methodology is presented. In section 4, an industrial case study is demonstrated, followed by a discussion in section 5, with section 6 providing the conclusions.

2. Preliminary concepts

2.1. Bayesian Network

The Bayesian network contains two parts: i) a graphical structure and ii) the conditional probabilities (1,2). The representation of the network structure is a directed acyclic graph (DAG) $G = (V, E)$, where V denotes the set of nodes and E denotes the set of edges of the graph structure. Each node $v_i \in V$ corresponds to an uncertain variable X_{v_i} in X_V for all i = 1,2, …, n. Each edge is a directed link between two nodes, which represents the direct causal relationship or the influence between linked nodes. The joint distribution P over the entire network can be factorized as:

$$P(X_V) = \prod_{i=1}^{n} P\left(X_{v_i}|X_{Pa(v_i)}\right)$$

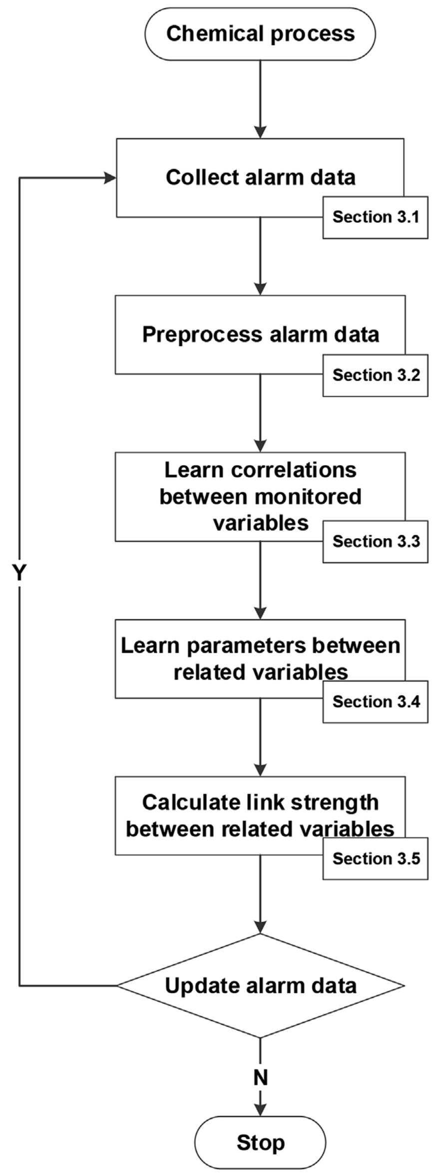

Fig. 1. The methodology of learning knowledge from alarm data.

where $X_{Pa(v_i)}$ denotes a set of parent variables of the variable X_{v_i}.

2.2. Structure learning

The structure is the qualitative information of the alarm data in Bayesian networks. The structure-learning problem is a Non-deterministic Polynomial-time hard (NP-hard) problem (Heckerman et al., 1995). This means it will become more and more challenging as the size of the learning data set and the number of variables increase. However, a properly designed search helps to find the most likely structure in a reasonable time.

Structure learning is an optimization problem which is comprised of two parts: i) score function and ii) search strategy. Given the specific scoring function as a criterion, the search strategy is employed to search for the structure which fits the dataset best.

Many score functions (Carvalho, 2009) measure the fitness between the dataset and the structure of variables. These scoring functions are mainly based on three different principles: i) entropy and information (Herskovits and Cooper, 1991; Chow and Liu, 1968), ii) the minimum description length (Bouckaert, 1993; Friedman and Goldszmidt, 1996; Lam and Bacchus, 1994; Suzuki, 1993), and iii) Bayesian approaches