Contents lists available at ScienceDirect

Process Safety and Environmental Protection

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

Risk assessment of rare events

Ming Yang[a,*], Faisal Khan[a], Leonard Lye[a], Paul Amyotte[b]

[a] Safety and Risk Engineering Group, Faculty of Engineering and Applied Science, Memorial University, St. John's, NL, Canada A1B 3X5

[b] Department of Process Engineering and Applied Science, Dalhousie University, Halifax, NS, Canada B3J 2X4

ARTICLE INFO

Article history:
Received 12 March 2015
Received in revised form 19 June 2015
Accepted 2 July 2015
Available online 14 July 2015

Keywords:
Rare event
Precursor
Hierarchical Bayesian Approach
Mutual information
probabilistic risk assessment
Bayesian network

ABSTRACT

Rare events often result in large impacts and are hard to predict. Risk analysis of such events is a challenging task because there are few directly relevant data to form a basis for probabilistic risk assessment. Due to the scarcity of data, the probability estimation of a rare event often uses precursor data. Precursor-based methods have been widely used in probability estimation of rare events. However, few attempts have been made to estimate consequences of rare events using their precursors. This paper proposes a holistic precursor-based risk assessment framework for rare events. The Hierarchical Bayesian Approach (HBA) using hyper-priors to represent prior parameters is applied to probability estimation in the proposed framework. Accident precursor data are utilized from an information theory perspective to seek the most informative precursor upon which the consequence of a rare event is estimated. Combining the estimated probability and consequence gives a reasonable assessment of risk. The assessed risk is updated as new information becomes available to produce a dynamic risk profile. The applicability of the methodology is tested through a case study of an offshore blowout accident. The proposed framework provides a rational way to develop the dynamic risk profile of a rare event for its prevention and control.

1. Introduction

Rare events are highly improbable undesired events but with severe consequences. Taleb (2007) has classified rare events into Black Swan and Gray Swan events. The former are unforeseeable and catastrophic events (e.g., the September 11 terrorist attacks) while the latter (e.g., offshore blowout accidents) can be potentially predicted yet not precisely. In this paper, "rare events" specifically refer to Gray Swan events. Historical information often indicates no or few occurrences of rare events. Thus, conventional statistical methods would produce biased and inconsistent estimates of the probabilities of rare events (Quigley and Revie, 2011; Quigley et al., 2007). The use of precursor data was proposed as one way to deal with the problem of data scarceness (Bier and Mosleh, 1990; Yi and Bier, 1998). Precursors are events that are moderate in severity and could have evolved to accidents (rare events) given further failures of safety barriers (Eckle and Burgherr, 2013; Johnson and Rasmuson, 1996). Meel and Seider (2006) have referred to precursors as near-misses and incidents, with incidents having higher probabilities of causing future accidents (i.e., rare events). They have also claimed that near-misses can indicate the likelihood of occurrence of future incidents and accidents and are helpful to prevent accidents. Khakzad et al. (2014) have recognized the usefulness of precursor data in rare event probability estimation from two perspectives:

(1) Precursors are reasonable indicators of rare events.
(2) Precursor data can be used as a foundation to construct likelihood functions in Bayesian approaches.

Bayesian methods are commonly used in precursor-based approaches for probability estimation of rare events

* *Corresponding author*. Tel.: +1 7098648126.
E-mail address: myang@mun.ca (M. Yang).
http://dx.doi.org/10.1016/j.psep.2015.07.004

(Kalantarnia et al., 2010; Khakzad et al., 2012; Meel and Seider, 2006). These approaches often model a rare event as a logical and sequential combination of system component and safety barrier failures through fault trees, event trees, bow-ties, and Bayesian networks. To capture the accident scenarios, some hazard identification techniques such as DyPASI (Paltrinieri et al., 2013) and ARAMIS (Delvosalle et al., 2006) are used. The determination of priors is a key issue in the Bayesian approaches. Posterior distributions are also highly sensitive to the choice of prior parameters due to scarcity of data. Furthermore, it has been shown that the Hierarchical Bayesian Approach (HBA) is capable of producing more consistent estimations than the conventional Bayesian approach (Chen and McGee, 2008; Yang et al., 2013).

Most literature focuses on the probability estimation of rare events using precursor data. However, to the authors' knowledge, limited work has been done on consequence estimation based on precursors. Losses caused by precursors should convey a certain degree of information associated with the potential consequence of a rare event. A sound estimation can be obtained based on the measurement of how much information a rare event and its precursor share with each other. Mutual information, a concept developed by Shannon (1948), provides such a measurement.

This paper aims to develop a precursor-based risk assessment framework for rare events. In this framework, the HBA is applied for probability estimation while mutual information is used to develop an estimate for the consequence. The rest of this paper is organized as follows. Section 2 presents the basics of the HBA. The concept of mutual information is discussed in Section 3. Section 4 explains the precursor-based framework for risk assessment of rare events. The application of the proposed method is demonstrated through a case study of an offshore blowout accident in Section 5. Finally, Section 6 provides the conclusions and the directions for future work.

2. Hierarchical Bayesian Approach (HBA)

A controversial part of any Bayesian method is the development of appropriate prior distributions (Siu and Kelly, 1998). The two-stage Bayesian method was first proposed by Kaplan (1983) to address source-to-source variability of data when incorporating available information to develop informative prior distributions. Kelly and Smith (2009) have claimed that the two-stage Bayesian model can be viewed as a particular example of a general hierarchical Bayesian model. The prior distribution for the parameter of interest, denoted $\pi(\theta)$ for a two-stage Bayesian model, can be written as (Kelly and Smith, 2009):

$$\pi(\theta) = \int \pi_1(\theta|\varphi)\pi_2(\varphi)d\varphi \quad (1)$$

where $\pi_1(\theta|\varphi)$ is the first-stage prior representing the population variability in θ; $\pi_2(\varphi)$ is the hyper-prior representing the uncertainty in φ; φ is a vector of hyper-parameters, e.g., $\varphi = (\alpha, \beta)^T$.

The prior is developed in multiple stages using generic data collected from different sources (e.g., different industrial sectors) in a general hierarchical Bayesian model. $\pi(\theta)$ is then used as an informative prior distribution to generate case-specific posterior distribution using case specific data.

The main advantages of the HBA over conventional Bayesian methods are its capability to model the population variability of data from different sources (Siu and Kelly, 1998) and to borrow strength from other indirect but relevant data (Yan and Haimes, 2010). Recent years have seen a few applications of the HBA to the probability estimation of rare events using precursor data (Khakzad et al., 2015; Khakzad et al., 2014; Yang et al., 2013). The hierarchical Bayesian model in this paper was developed based on the above studies.

3. Mutual information

Mutual information is a concept from information theory developed in the context of digital communication (Shannon and Weaver, 1949). Mutual information can either be interpreted as the distance from independence between two random variables (Cover and Thomas, 1991) or as reduction of uncertainty of one random variable given the knowledge of the other one (DeGroot, 1962). The latter is the definition used in this paper and is further discussed in the following paragraphs.

Shannon (1948) has used the concept of entropy as a measure of uncertainty of a random variable. Let X be a discrete random variable with probability mass function $P_x(x)$. The entropy $H(X)$ is defined as (Lu, 2011):

$$H(X) = -\sum_{x\in X} P_x(x)\log P_x(x) \quad (2)$$

where the log is to the base of 2 and the unit of $H(X)$ is bit.

The mutual information between two random variables (i.e., X and Y) is defined as follows (Cover and Thomas, 1991):

$$I(X;Y) = \sum_{x\in X}\sum_{y\in Y} P_{x,y}(x,y)\log\frac{P_{x,y}(x,y)}{P(x)P(y)} \quad (3)$$

where $I(X;Y)$ is the mutual information between X and Y; $P_{x,y}(x, y)$ is the joint probability mass function; $P(x)$ and $P(y)$ are marginal probability mass functions.

The above equation can be further rewritten to show the relationship between mutual information and entropy:

$$\begin{aligned} I(X;Y) &= -\sum_{x\in X} P_x(x)\log P_x(x) - \left[-\sum_{x\in X}\sum_{y\in Y} P_{x,y}(x,y)\log P_{x|y}(x|y)\right] \\ &= H(X) - H(X|Y) \end{aligned} \quad (4)$$

where $P_{x|y}(x|y)$ is the conditional probability mass function; $H(X|Y)$ is the conditional entropy of X given Y.

Sarndal (1974) first introduced a normalized version of $I(X;Y)$ to represent the relative reduction of uncertainty contained in X given the knowledge of Y:

$$U(X;Y) = \frac{I(X;Y)}{H(X)} \quad (5)$$

where $U(X;Y)$ is known as the asymmetric uncertainty coefficient.

Sarndal (1974) has also proposed a symmetric version of $U(X;Y)$:

$$S(X;Y) = \frac{I(X;Y)}{1/2[H(X) + H(Y)]} \quad (6)$$

The concept of mutual information has been widely used for system feature selection (Hoque et al., 2014), system