Process Safety and Environmental Protection xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Process Safety and Environmental Protection

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

Simulation-based fault propagation analysis—Application on hydrogen production plant

Hossam A. Gabbar[a,b,*], Sajid Hussain[a], Amir Hossein Hosseini[a]

[a] Faculty of Engineering and Applied Science, University of Ontario Institute of Technology, 2000 Simcoe Street North Oshawa, Ontario L1H7K4, Canada
[b] Faculty of Energy and Nuclear Science, University of Ontario Institute of Technology, 2000 Simcoe Street North Oshawa, Ontario L1H7K4, Canada

ABSTRACT

Recently production of hydrogen from water through the Cu–Cl thermochemical cycle is developed as a new technology. The main advantages of this technology over existing ones are higher efficiency, lower costs, lower environmental impact and reduced greenhouse gas emissions. Considering these advantages, the usage of this technology in new industries such as nuclear and oil is increasingly developed. Due to hazards involved in hydrogen production, design and implementation of hydrogen plants require provisions for safety, reliability and risk assessment. However, very little research is done from safety point of view. This paper introduces fault semantic network (FSN) as a novel method for fault diagnosis and fault propagation analysis by using evolutionary techniques like genetic programming (GP) and neural networks (NN), to uncover process variables' interactions. The effectiveness, feasibility and robustness of the proposed method are demonstrated on simulated data obtained from the simulation of hydrogen production process in Aspen HYSYS®. The proposed method has successfully achieved reasonable detection and prediction of non-linear interaction patterns among process variables.

Keywords: Fault sematic network (FSN); Cu–Cl thermochemical cycle; Aspen HYSYS®; Genetic programming; Neural networks; Process variables interaction

1. Introduction

Inferring the interaction structures among different process variables (PVs) from their observed dynamics is one of the important tasks in predicting fault propagation behaviors of complex chemical, petrochemical or nuclear plants. To achieve the above goals, PVs emanating from chemical, petrochemical or nuclear plants are recorded and analyzed. Analysis of these variables is critical and challenging for data analysis community and such a problem is covered by system identification theory. In a plant operation, there could be serious consequences because of human error or any plant malfunction including equipment or component failure. Also the overall plant availability may reduce because of the force shutdown or the final product quality may suffer. Catastrophic events may also occur due to release of toxic chemical or fire outbreak. In usual case, abnormal deviation of functional process variables or key performance indicators (KPIs) triggers an alarm in the control room and the process operator has to take necessary and timely remedial actions. In order to do that, the operator must know the exact cause and consequence of the deviations, their relationships and propagation behaviors.

Hydrogen is currently gaining much attention as a new energy source which can be replaced with oil and other fossil fuels in near future not only in different industries but also in transportation sector. Like other industrial process, hydrogen production process is also involved with different hazards that may lead to irreparable losses that not only affect the

* Corresponding author at: Faculty of Engineering and Applied Science, University of Ontario Institute of Technology, 2000 Simcoe Street North Oshawa, Ontario L1H7K4, Canada. Tel.: +1 905 721 8668x5497; fax: +1 905 721 3046.
E-mail address: hossam.gabar@uoit.ca (H.A. Gabbar).

Received 8 October 2013; Received in revised form 4 December 2013; Accepted 20 December 2013

http://dx.doi.org/10.1016/j.psep.2013.12.006

equipment and production but raises environmental, occupational safety and health related concerns. Given this it is obvious that performing preventive actions and implementing well-established monitoring and control strategies to prevent accidents and reducing the risk associated with process seems necessary.

Zalosh et al. made a statistical investigation in 1978 about industry related hydrogen accidents in USA and he concluded that in about 80% of the hydrogen accidents ignition occurred and hereof about 65% of these ignitions caused an explosion. Forty percent of all the hydrogen leakages were not detected prior to the accident and therefore it was argued to install appropriate detectors in hydrogen systems (Markert et al., 2007). There are many ways that make hydrogen different from conventional fuels such as representing a greater hazard over methane and gasoline due to the wider flammability limits, lower ignition energy and higher deflagration index. Consequently, greater risk due to the increased probability of a fire and explosion would be one of the results. Given this, it seems necessary to perform a detailed hazard identification method for every stage in the hydrogen supply chain.

This paper proposes a systematic hazard identification and risk assessment method based on FSN. The rest of the paper is organized as follows: Section 2 describes the theory behind FSN followed by Section 3 that describes the basics of GP and Section 4 that discusses non-linear autoregressive model with exogenous inputs (NARX). Section 5 formulates the objective function followed by hydrogen process simulation in Section 6. Section 7 presents the simulations and results followed by Section 8 that discusses the practical applications of FSN and conclusion is drawn in Section 9.

2. Fault semantic network (FSN)

The concept of semantic network was first proposed by Richens (1956). It is a network structure that represents relations between concepts. The concept of semantic network was further developed by Collins and Quillian (1969) where they introduced semantic network in a tree structure (directed or undirected graph) consisting of nodes and arcs. The nodes represent concepts and the connections show relations between nodes. The FSN was originally realized by Gabbar (2007) and further elaborated by Gabbar (2010) and Gabbar and Khan (2010) is a mean of representing fault knowledge based on relationships between objects. In FSN, the nodes correspond to different faults/causes/consequences and the links between them describe the dependencies. Initially, FSN is constructed based on ontology structure of fault models on the basis of process object oriented methodology (POOM) where failure mode (FM) is described using symptoms, enablers, variables, causes, consequences, and repair actions. Rules are associated with each transition of the causation model within FSN. The rules can be quantitative (probabilistic) or qualitative. For example, failures related to gear tooth breakage might be associated with a qualitative rule such as

IF Structure = Gears AND PV = Vibration AND Symptom = Mesh Frequency Sidebands AND Dev = Very-High THEN FM = Gear Tooth Breakage AND Consequence = Damage/Production Loss AND Repair = Replacement.

These rules are initially defined in generic form based on domain knowledge, i.e., regardless of plant specific knowledge and then further explained or trained for plant specific knowledge based on observations. As described above, the structure of FSN represents relations between variables quantitatively or qualitatively and this can be done through (1) probabilistic approach, (2) fuzzy approach and (3) mathematical model approach. Each approach is discussed as follows:

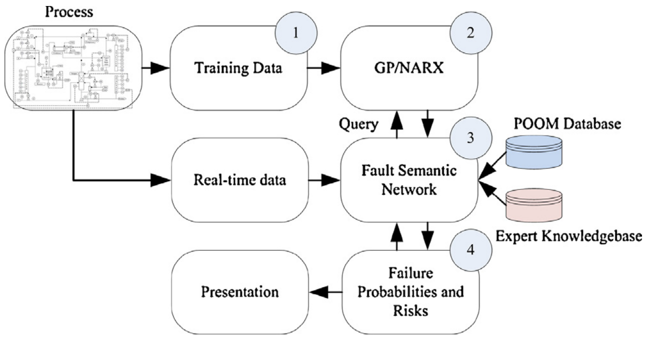

Fig. 1 – Process to construct the FSN.

Probabilistic approach: In probabilistic approach, a probability value is assigned to each node depending upon its hierarchy in the network as a parent node or a child node. Bayesian belief networks (BBN) are used in this approach.

Fuzzy approach: This type of reasoning is a pure qualitative reasoning. In fuzzy approach, rules are associated with each transition of the causation model within FSN. In other word, relations between variables are described by specifying if-then statements.

Model formulation: This type of reasoning is a pure quantitative reasoning. Relationship between two variables is specified by mathematical equation such as $y = \left[\log(x + \cos x) - \cos^2 x\right]$, where x is an independent variable and y is a dependent variable. The mathematical models can be derived from system identification theory using genetic programming (GP), neural networks (NN), adaptive neuro-fuzzy inference system (ANFIS) or other statistical based methods.

The process of constructing FSN is shown in Fig. 1, where training data are recorded from the real world or simulated process.

After real world data collection or process simulation, PVs interaction learning framework based on GP follows to uncover the complex PVs interaction patterns. The learned GP mathematical models are then used for constructing FSN. After the system has learned the dynamics of the PVs and has uncovered the interaction strengths among PVs, the system can be deployed on real-time process data. In real-time deployment, the PVs are monitored and the state observation for each PV is passed to the FSN. The FSN is responsible to interpret the results in order to calculate the failure probabilities and associated risks. The query process also updates the FSN database or GP mathematical models if new evidence is encountered. The ultimate goal of the study is to develop a real-time fault propagation analysis and hazard identification method through FSN. The overall process can be described in three steps as follows:

- *Extracting real-time process data*: The real-time data are extracted from sensors and controllers installed on all equipment in the underlying plant to monitor the process conditions.
- *Relationship among process variables through GP*: In order to uncover the relationship among process variables, GP is used as pattern recognition techniques to identify their relationship quantitatively. This step is critical in order to analyze propagation of faults among process variables.