Risk Analysis using Large Alarm Databases

Yilu Fan
Department of Chemical and Biomolecular Engineering, University of Pennsylvania
Philadelphia, PA 19104 USA (E-mail: yilufan@seas.upenn.edu)

Warren D. Seider
Department of Chemical and Biomolecular Engineering, University of Pennsylvania
Philadelphia, PA 19104 USA (Tel: 215-898-7953; e-mail: seider@seas.upenn.edu)

Ulku G. Oktem
Risk Management and Decision Processes Center, Wharton School
Philadelphia, PA 19104 USA (Tel: 215-573-7704; e-mail: oktem@wharton.upenn.edu)

Ankur Pariyani
Near-Miss Management LLC
Philadelphia, PA 19106 USA
(e-mail: pariyani@nearmissmgmt.com)

Masoud Soroush
Department of Chemical and Biological Engineering, Drexel University
Philadelphia, PA 19104 USA (Tel: 215-895-1710; e-mail: masoud.soroush@drexel.edu)

Jeffrey Arbogast
American Air Liquide
Newark, DE 19701 USA (Tel: 302-286-5437; e-mail: jeffrey.arbogast@airliquide.com)

Darrin Feather
American Air Liquide
Newark, DE 19701 USA (Tel: 302-286-5469; e-mail: darrin.feather@airliquide.com)

In this work, our methods for the dynamic risk analysis of large alarm databases [1-3], previously tested on a fluidized catalytic cracking unit (FCCU), are extended and tested using data over three years from a steam-methane reformer (SMR) unit at Air Liquide. This paper also introduces methods of providing online estimates that alert operators to potentially serious safety problems. The estimates permit the operators to make adjustments that avoid trips and accidents.

The SMR unit has four safety systems, a basic process control system, first- and second-layer alarms, and an automatic emergency shutdown system. The success or failure of each safety system leads to different end-states of the operating unit, either continued operation (CO) or emergency shutdown (ESD). To estimate the failure probabilities of each safety system, 2.7 million DCS and ESD alarm data records are compacted. For each alarm entry, its module, the time at which the alarm was triggered, the alarm category, the alarm state, and the alarm level were recorded.

For data compaction analysis, the raw data are filtered. MATLAB programs count the number of the most-critical, moderately-critical, and least-critical abnormal events. First, the “module” column is sorted to group together alarms having similar “States” or “Levels”. For the large number of abnormal events associated with each module, the starting and ending times of each abnormal event are labeled by the “State” information for its module. The MATLAB codes, operating on a large alarm matrix, containing 1,152,980 rows of alarm entries, find 5,113 most-critical abnormal events, 296 moderately-critical abnormal events, and 70,240 least-critical abnormal events. In addition to the alarm database, Air Liquide provided information about the plant trips for the SMR unit over the study period.

The paths through event trees for the safety systems show that nearly all abnormal events, evidenced by alarms associated with process variables, are identified as their variables are returned to normal operation. This vast amount of "near-miss" data improves our estimates of the failure probabilities of the safety systems. The success and failure counts for each safety system are input to Bayesian analysis to estimate the failure probabilities and the probabilities of trips and accidents. To perform the Bayesian analyses, the random-walk, multiple-block, Metropolis-Hasting (H-M) algorithm is used. The algorithm draws samples from approximate distributions, known as proposal or jumping distributions, and corrects them to better estimate the target posterior distribution for each safety system. More specifically, proposal distributions are chosen to be random-walk distributions. For high-dimensional systems, such as the SMR unit, a single block M-H algorithm that converges rapidly to the target density can be difficult to construct. In such cases, smaller blocks are obtained by discretizing the multivariate space, and Markov chains are constructed in the smaller blocks. Sequentially, a single iteration of the multiple-block, M-H algorithm is completed by updating each block in sequence. Many iterations are implemented until the initial state is indiscernible.

Once the failure probability distribution, the mean failure probabilities, and the probabilities of trips and accidents are obtained, process data that are less discrete than alarm data are utilized, especially taken during periods of "alarm flooding", to improve the probability estimates. As in our earlier analyses [2-3], copulas represent the interactions between the failure probabilities of the safety systems. Because the failure probabilities are not independent, with their marginal distributions belonging to different families, copulas are used to describe their joint distribution, marginals are combined, and interactions between the random variables are accounted for using the dependence matrix. The dependence matrix also enables the random variables to interact with each other and share information. This is especially important when limited data are available for certain random variables. Thus, because of the ability to model arbitrary pairwise correlations between the random variables, the multivariate normal and Cuadras-Auge copulas are used. In particular, the performances of these two copulas for SMU unit are compared with those for the FCCU.

References

1. Pariyani, A., W. D. Seider, U. G. Oktem, and M. Soroush, "Incidents Investigation and Dynamic Analysis of Large Alarm Databases in Chemical Plants: A Fluidized-Catalytic-Cracking Unit Case Study," *Ind. Eng. Chem. Res.*, **49**, 8062-8079 (2010).

2. Pariyani, A., W. D. Seider, U. G. Oktem, and M. Soroush, "Dynamic Risk Analysis using Alarm Databases to Improve Safety and Quality: Part I – Data Compaction," *AIChE J.*, Early View.

3. Pariyani, A., W. D. Seider, U. G. Oktem, and M. Soroush, "Dynamic Risk Analysis using Alarm Databases to Improve Safety and Quality: Part II – Bayesian Analysis," *AIChE J.*, Early View.