Second Annual Symposium, Mary Kay O'Connor Process Safety Center
"Beyond Regulatory Compliance: Making Safety Second Nature"
Reed Arena, Texas A&M University, College Station, Texas
October 30-31, 2001

Development of a Decision Support System for Chemical Accident

Gaurav Sharma,
W. Rogers, and M.S. Mannan
Mary Kay O'Connor Process Safety Center
Chemical Engineering Department Texas A&M University
College Station TX 778433122

Phone: 979/862-3985
Email: gausha@hotmail.com

ABSTRACT

Decision support systems (DSS) are used extensively in business to enable managers to make future decisions. Every business area requires a customized DSS. The uniqueness of each business area stems from various types of data requirements for decision-making processes. Therefore, the developer must analyze each industrial area for its particular requirements before designing a DSS.

The back-end of a decision support system involves comprehensive modeling of the issues related to indexing and storing data. The data in the back-end are subjected to data-mining procedures, which form the engine for retrieval of relevant and accurate information to the decision-maker. These procedures form the basis for Knowledge Discovery in Databases (KDD). The front end of a DSS involves statistical and mathematical models to convert the data into information for the manager. The statistical functions to be employed for processing data depend upon the requirements of the decision-maker. These issues applied to the chemical accident data demand an innovative out-of-box methodology.

This paper explores the development of an appropriate decision support system for chemical accident data. Taxonomy for indexing accident data and the design of a corresponding database model are established. Innovative data mining techniques are also explored. Finally, mathematical and statistical models are introduced for more effective processing of accident data.

Development of a Decision Support System for Chemical Accident Information

Gaurav Sharma and M.S. Mannan
Mary Kay O' Connor Process Safety Center
Chemical Engineering Department
Texas A&M University
College Station TX 77843-3122

Abstract

This paper discusses the functionalities of the Decision Support System (DSS) for chemical accident information being developed at the Mary Kay O'Connor Process Safety Center (MKOPSC). The relational database model has been used to relate accident information with chemicals used, equipment employed, and industry location. It is proposed that the DSS provide ready-to-use information to industry personnel so that they are acquainted with the risks involved in plant operations. Also, statistical analysis has been proposed to compare the severity of accident consequences due to various chemicals. Such analysis would help to differentiate high-risk chemicals from the low-risk ones. Similar statistical analysis has been proposed for comparing various equipment that underwent failure during the chemical accidents. The purpose of this approach is to inform industry personnel about the accident consequences that occurred in the past involving equipment. Lastly, the paper proposes that the DSS have the functionality to study the effect of accident location on accident consequences. Thus, at the heart of the DSS is a relational model between accident consequences and the factors that influence those consequences.

Introduction

Decision Support Systems (DSS) are used extensively in business to enable managers to make future decisions [4]. The purpose of a DSS is to use archived data and present it in a way that influences the decision-making duties of a manager [4]. In the past, process safety applications of such systems have been lacking. The aim of this paper is to describe the functionalities of a DSS being developed at MKOPSC to help chemical industry personnel make better safety decisions.

Statistical Analysis

Whenever chemical accidents occur the consequences can be gauged in terms of monetary loss and the cost to human life [1]. Both these consequences are grave and we do not wish to choose either over the other to determine the severity of the accident. Therefore, in our statistical analysis we will capture both these factors. Normal distribution for the random variable will be assumed in all statistical analysis proposed for the DSS. Normal distribution, which is one of the most commonly used distribution in statistical analysis, was originally derived by De Moivre in 1733 [2]. We wish to apply this distribution to analysis of chemical accident data. In order to do so, we will determine the probability density function from the mean and variance of the random variable. Statistical analysis for calculating probabilities using this probability density function can be performed with statistical computational tools such as SPSS and SAS [2].

Storing and Indexing Data

The first step for the development of a database is preparing a model for storing the data. The database modeling should reflect the actual industry function that the database application is automating [2]. In our case it is not a business process that we are trying to computerize, but we wish to use the chemical accident information for making intelligent safety decisions. Therefore, the database model must encapsulate the various entities involved in chemical accidents, whether they are the consequences or the factors influencing the consequences. Since, these two are closely related, the relational database model must reflect this relationship. It is this relationship between causes and consequences that form the framework of the DSS.

There are best practices for modeling data in the form of Entity Relationship Diagrams (ERD) [3], which are road maps for database application developers for defining tables and queries [3]. Also, they serve as pointers for future developers who want to use the application, whether for upgrade or for merging with other database applications. Another best practice in database development is that of normalization, which ensures efficient usage of storage space, reduces redundancy of data, and prevents anomalies from occurring in the database model [3]. Both of these practices have been applied in the database model for the DSS being developed at MKOPSC.

There are four entities in the database namely accident, location, chemical, and equipment, which are stored in separate tables. The relationship among these entities is described in Figure 1.

The Chemicals Table

The chemicals table contains physical and chemical properties for all possible chemicals. The purpose of this table is to provide ready-to-use information for industry personnel before dealing with any chemical. There is a search facility that an operator can use to search for a particular chemical he is going to work with. Attributes like <Effect on Human Body> and <Precaution to be Taken> give forewarning to the operator what to expect. Also, a list of chemicals that react explosively with the particular chemical is provided. The idea of this information is to make sure that anybody storing, processing, or designing a chemical process with that particular chemical must be equipped with enough information about that chemical so that accidents are prevented.

Statistical Analysis of Chemicals Involved in Accidents

The purpose of statistical analysis of various chemicals involved in accidents is to distinguish between chemicals causing greater risk to human life and those posing mild risk. The basis of such statistical analysis is to assume that the severity of an accident is based on the number of fatalities it causes. Since the chemicals table has a relational model with the accident table, we can extract the number of fatalities for various accidents that have occurred involving that chemical. These numbers can be assumed to fit a certain distribution, and we assume that the number of fatalities for a particular accident follows a normal distribution. This assumption can be verified using statistical computational tools like SPSS. On that assumption, we can calculate the probability of greater then 'x' number of fatalities for each chemical. The greater the probability, the greater is the risk involving that chemical. This is the nature of statistical analysis for the chemicals table.

Next we shift attention to the economic consequences of chemical accidents. We wish to calculate the probability of dollar damage greater than 'x' for all possible chemicals. This allows comparison of dollar damage probabilities for various chemicals. Similar to the analysis for fatalities, this analysis helps us to differentiate chemicals causing low monetary loss to those causing greater economic consequences.

The Equipment Table

The equipment table contains fields for storing data about all possible equipment used in the chemical industry. Each unit is characterized by attributes like <Type of Equipment>, <Manufacturer Name>, and <Age>. There is a relational model between the equipment table and the accident table. It implies that we are able to query the accident table for incidents involving particular equipment, which enables comparison of accident frequencies among different equipment. Such a comparison provides information to industry personnel about the equipment, which upon failure cause greater loss to human life. This knowledge can be used to prevent high loss to human life during accidents.

Another conclusion that can be drawn from the equipment-accident relational model is whether equipment manufactured by a particular manufacturer is more susceptible to failure. This information can assist the procurement department in making future purchase decisions.

Age of equipment can be another factor causing accidents, so we can use the equipment table to determine the age after which a particular type of equipment exhibits a tendency to fail. This information can serve as a guide to the procurement department in replacing equipment.

Statistical Analysis of Equipment Data

Just like in the case of chemicals, we assume that the severity of an accident is determined by the number of fatalities it causes and that the frequency of fatalities follows a normal distribution. Using these assumptions we calculate the probability of greater than 'x' number of fatalities for all possible types of equipment. This kind of statistical analysis informs the industry personnel if certain types of equipment result in more severe accident consequences than others. This can prove to be useful when safety policies for the plant are being developed. Similarly, using the accident data for various manufacturers, we calculate the probability of greater than 'x' fatalities for all possible equipment. This analysis can give a comparison of which manufacturers' equipment results in greater risk. As mentioned above, this comparison provides decision support to the procurement department in their purchases.

The Equipment-Chemical Relational Model

So far we have stressed the relational model between equipment-accident and chemical-accident. Another important relational model that can be used is that between the equipment and chemicals. The aim is to study whether certain chemical-equipment combinations occurred more frequently during chemical accidents. For example, let us take the equipment 'x'. If a large number of accidents involving equipment 'x' also involve chemical 'y', this combination of equipment-chemical can become a point of investigation. If a trend involving equipment 'x' and chemical 'y' is discovered, then preventive safety measures can be implemented.

Location-Accident Relational Model

The aim of the location-accident relational model is to determine whether location of the accident site can be correlated to the severity of the accident. In this case, severity is defined by the number of fatalities caused by the accident. The number of fatalities is divided into two groups namely 1. On-site fatalities and 2. Off-site fatalities. The <Distance of the Chemical Plant from the nearest city> is one of the attributes used to study the effect of location on the accident severity. Greater distance from the city can imply that emergency medical care will be delayed. At the same time, the off-site fatalities may be reduced due to fewer people living in the area close to the industry. It is proposed that this trend can be studied using the DSS.

Conclusions

The purpose of the DSS being developed at MKOPSC is to provide processed chemical accident information to industry managers. The central point of the DSS is the relationship between the accident consequences and factors, namely chemicals, equipment, and location, which affect the consequences. This relationship can be extended to other factors that affect the consequences of the chemical accident, which will be a part of the subsequent research.

References

1. Helena Montiel, Juan A. Vílchez, Josep Arnaldos and Joaquim Casal, "Historical analysis of accidents in the transportation of natural gas," *Journal of Hazardous Materials, Volume 51, Issues 1-3, November 1996, pp. 77-92*
2. Douglas C. Montgomery, George C. Runger, "Applied Statistics and Probability for Engineers," John Wiley & Sons Inc 2nd Edition
3. Candace C. Fleming, Barbara von Halle, "Handbook of Relational Database Design," Addison-Wesley Publishing Company
4. Ian T. Cameron, "An interactive web-based decision support system for hazardous industry land-use planning," *Computers & Chemical Engineering, Volume 24, Issues 2-7, 15 July 2000, pp. 1057-1062*

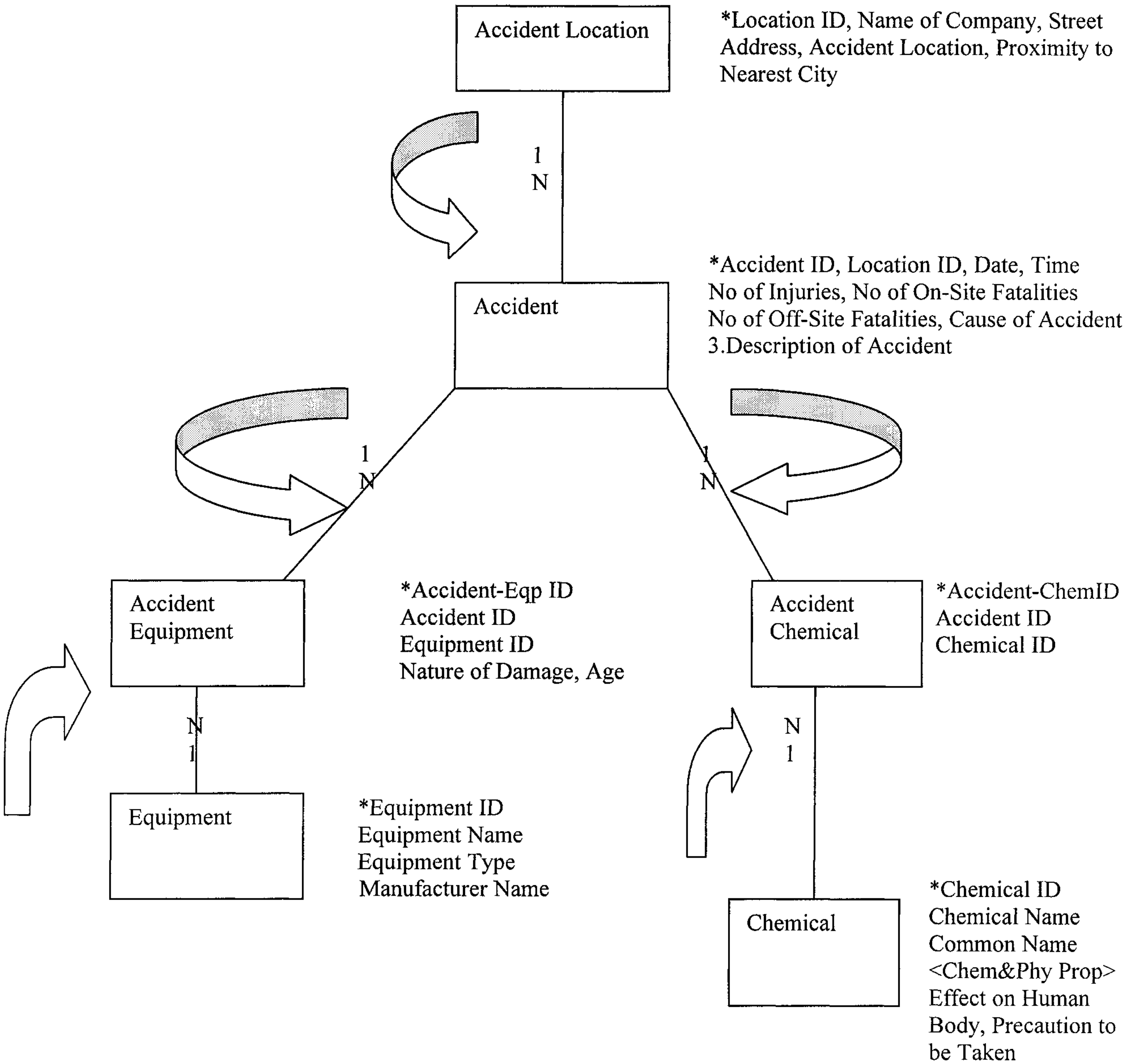

Figure 1. Database Model for Chemical Accident Information