

A Study on Developing IGAS for Analyzing the Status of Gas Facility

Jeong Seok Oh

*Institute of Gas Safety R&D, Korea Gas Safety Corporation, 1390, Wonjung-ro,
Maengdong-myeon, Emseong-gun, Chungcheongbuk-do, Korea
jsoh90@gmail.com*

Abstract

Gas facilities such as LPG station have showed frequently diverse accidents, because they are seemed to satisfy only institutional requirements. In Gas accidents, human is damaged directly because most cause is careless handling and inadequate safety equipment, most type is explosion, fire and rupture. Therefore, an advanced safety process is required for enforcing gas safety management. Although correct progress direction in advances safety process is necessary for many devices, methods and system, it is to require analysis method that many different data are analyzed in parallel but each measuring data is analyzed in individual. This paper preferentially aims to check the feasibility of machine learning analysis in order to apply a safety of gas facility, and devise method using artificial intelligent algorithm in order to manage total safety analysis that can analyze simultaneously many different data. At First, the feasibility study must be generally selected target gas facility, collected what kinds of risk factor, and then considered the appropriated machine learning method. Next, our research develops total risk analysis algorithm with a combination method between classification and clustering algorithm. Finally, we developed IGAS (Intelligent Gas Analysis System) for small gas facility version. This method and system are to mark the beginning of analysis method for detecting cause and increasing safety about gas facilities.

Keywords: *the feasibility check, machine learning, gas safety domain*

1. Introduction

Safety management method must be considered as all energy industries because of their own inherent risks. The safety management for gas facilities should be considered more carefully because those facilities have been operated by various kinds of materials. Especially, LPG station, LPG store and user less than medium-size have showed frequently diverse accidents, because they are seemed to satisfy only institutional requirements. In gas accidents, human is damaged directly because most cause is careless handling and inadequate safety equipment, and most type is explosion, fire and rupture. Therefore, an advanced safety process is required for enforcing gas safety management. Although correct progress direction in advances safety process is necessary for many devices, methods and system, it is to require analysis method that many different data are analyzed in parallel but each measuring data is analyzed in individual. So, deriving risk analysis is very important for preventing and corresponding accidents by means of a specific analysis method.

Recently, machine leaning methods are tried to use in this gas safety domain in order to analyze risk and accident sign. Data mining algorithms in machine learning have been used only without the reliability check. So, most attempts to apply gas facilities are unavailing until now. Otherwise, we cannot understand whether machine learning method is applied properly in gas safety domain or not. Furthermore, core logics of machining learning methods are modified or refined for adjusting gas safety domain in order to use

properly those methods. In preceding paper [5, 6, 9], we checked the feasibility about applying artificial intelligence algorithm to gas domain, and suggest methodology for enhancing accuracy.

This paper preferentially aims to check the feasibility of machine learning analysis in order to apply a safety of gas facility, and devise method using artificial intelligent algorithm in order to manage total safety analysis that can analyze simultaneously many different data. The feasibility study must be generally selected target gas facility, collected what kinds of risk factor, and then considered the appropriated machine learning method. Our study will want to know the status of facility using collecting various measure values. As a result, these machining learning algorithms are able to use in gas domain for enhancing safety of facilities. However, increasing accuracy of machine learning model is necessary to refine the nature of algorithm depending on gas domain characteristics. Furthermore, this research develops total risk analysis algorithm with a combination method between classification and clustering algorithm. This method is to mark the beginning of analysis method for detecting cause and increasing safety about gas facilities. So, we develop IGAS (Intelligent Gas Analysis System) for small gas facility version. The purpose of IGAS with our proposed algorithm must be complex algorithm whereby it is to minimize weakness and maximized advantages in gas safety environment.

2. Evaluating Feasibility using Intelligent Algorithm for Investing Data of Gas Facilities

Our first feasibility study would try to use real data and simple method. This examination selected pipeline as target facility, collected values of protective potential , and considered Bayesian algorithm as simple machine learning method in order to decide whether a gas facility seem to be corroded or not. Furthermore the status of protective potential is divided into normal, care, sign and warning like as usual status division method of protective potential in gas industrial.

Although “navie” Bayesian has been tested as efficient classification method, it does not apply to gas safety environment because it must read whole data every time probability is calculated without creating model. Because using machine learning model would be appropriate, our research is to create an inference model with Bayesian inference algorithm, refine row data, and execute status classification using thus model. Especially, to refine row data can include various operations on the basis of filtering degree. Our research collects about 10,000 real data in gas T/B boxes like as Table 1

Table 1. Table Label

NO	Measure time	Max(mV)	Min(mV)	Avg(mV)	Voltage
1	2014-08-16 14:00:00	-747	-1129	-942	3.9
2	2014-08-16 02:00:00	-1125	-1129	-1127	3.9
3	2014-08-15 14:00:00	-917	-1144	-1083	3.9
4	2014-08-15 02:00:00	-1139	-1153	-1083	3.9
5	2014-08-14 14:00:00	-802	-1159	-1075	3.9

Test box data are made up measure time, maximum potential, minimum potential, average potential, and battery voltage. Also real data range shows Figure 1. In Figure 1, the normal status is potential value (mV) between -2500mV and -850mV. The care status

is value between -850mV and -600mV, and the warning status is value less than -2500 or over than -100.

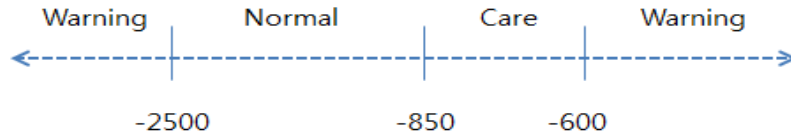


Figure 1. Real data range Example

Data input of Bayesian inference algorithm is normalized inner data scope for recognizing each status as pattern like as Figure 2. In Figure 2, the normal status is value between -2160 and -840. The care status is value between -2400 and -2160 or between -840 and -600. The sign status is value between -2700 and -2400 or -600 and -300. Finally, the warning status is value less than -2700 or between -300 and 0

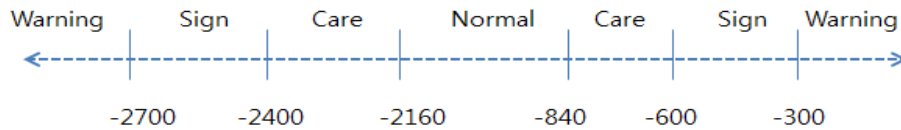


Figure 2. Normalized inner data scope

Also, likelihood of sufficiency (LS) and likelihood of necessity (LN) per rule are provided by experts in order to classify normal, sign, care, and warning pattern. The LS means reliability degree of hypothesis H that is treated by experts, if evidence E exists. The LN means unreliability degree of hypothesis H that is treated by experts, if evidence E does not exist. In our test, inference results using Bayesian algorithm show probability each four patterns like as Figure 3. The accuracy of results is about 80% by validation of experts. According to cases, this Bayesian inference model is necessary to optimize algorithm for gas safety domain in order to improve accuracy.

Figure 3. Classification process using Bayesian Algorithm

Furthermore, our study validates one more time in a different way for verifying efficient intelligence decision. Also, this study use decision tree that support WEKA

software. The WEKA software provides various machine learning algorithms as open software. Especially, this research use J48 algorithm that was modified C4.5 of decision tree algorithms in order to reduce error rate. The J48 algorithm can create a decision model as result, and this decision model is appropriate generally more than 90% accuracy for using application. The creation of decision model is based on potential protective data utilized when using Bayesian algorithm. This decision tree model is about 92.2% accuracy by 10-cross validation method. In this paper, more than 90% accuracy would mean usable model in this domain for future classification

The second feasibility study selected exposed pipeline as target facility, collected various values such as temperature, humidity, illuminance, infrared rays, and measuring point. The measuring point means class name, and is classified into underground, inner and outer. As a result, that model shows simple tree which can be classified into classes such as underground, inner and outer by only using infrared rays and humidity. In the concrete, this tee classified into underground with under 171V infrared rays and over 39% humidity, and outer with under 171V infrared rays and under 39% humidity. This model shows about 99% accuracy by 10-cross validation method. This model can use when new data want to know belonging to some environment, but it is inadequate for finding facility location

3. Applying Intelligent Algorithm for Gas Domain

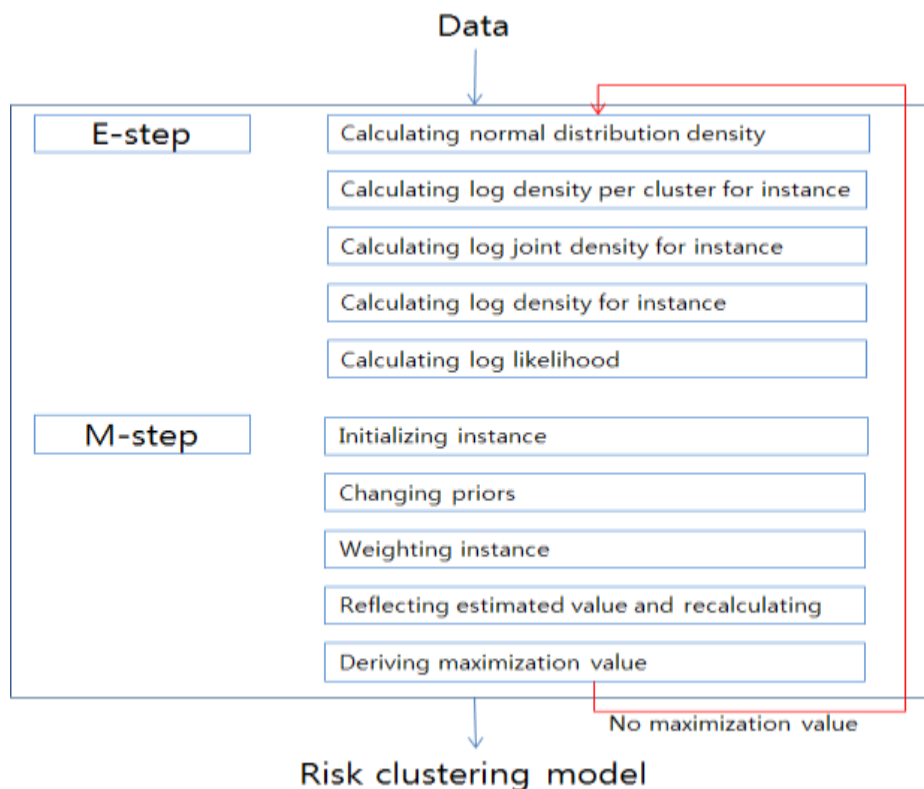


Figure 4. The Clustering flow with EM algorithm

The algorithm, which is necessary to generate total analysis model in gas supply facilities under medium-size, should be realized through improving itself and combining among various algorithm. Decision tree and clustering method is considered that t may be appropriate to algorithm for total risk analysis [1, 2, 3, 4]. Therefore, our research will modify and use EM (expectation-maximization) algorithm as clustering method. The EM algorithm is classified roughly into expectation procedure and maximization procedure.

The expectation procedure calculates expected value of log likelihood as estimated values at parameter. In other words, expectation procedure decides cluster of the highest probability that should belong to clustering by comparing previous selected log likelihood and current computed likelihood. The first step calculates cluster density through normal distribution density function.

$$F(x, m, \sigma) = \left(\frac{(x - m)^2}{2 \times \sigma^2} \right) - \log(\sqrt{2\pi}) - \log(\sigma) \quad (1)$$

The second step gets weighted value between instance and cluster and finds linking density about instance units

$$W(j) = \sum_{c=0}^j F(x(c), m(c), \sigma(c)) \quad (2)$$

$$W(j) = \sum_{c=0}^j F(x(c), m(c), \sigma(c)) \quad (3)$$

The final step works out density about whole instances and calculates log likelihood at estimation procedure

$$l(\sigma) = d(l) + \exp \sum_{\alpha=0}^l d(\alpha) - d(l) \quad (4)$$

$$\text{Log}lk = \frac{\sum_{c=0}^j lW(c) \times l(c)}{\sum_{c=0}^j lW(c)} \quad (5)$$

The maximization procedure generates variable value by maximizing expected value of the former procedure. The first step initializes instances and changes priors for reflecting next expectation procedure. The next step weights instance by changing mean and standard deviation estimated value about whole instance. The final step reflects and recalculates estimated value, and then derives maximization value.

$$\sigma^2(i, j) = \frac{\left(\frac{\sigma(i, j) - m(i, j)^2}{W(i, j)} \right)}{W(i, j)} \quad (6)$$

$$\sigma = \sqrt{\sigma^2(i, j)} \quad (7)$$

$$m(i, j) = \frac{m(i, j)}{lW(i, j)} \quad (8)$$

Also, our research selects decision tree for classifying status. A decision tree ramifies nodes by finding the best break point per each node in tree, and then branching node is repeated until node does not ramify any more to bottom nodes. Our algorithm use decision tree as prediction model linking between observed value and predictive value. A very critical matter in this tree structure should decide leaf nodes with no more of branching. In our research, classification procedures are shown Figure 5.

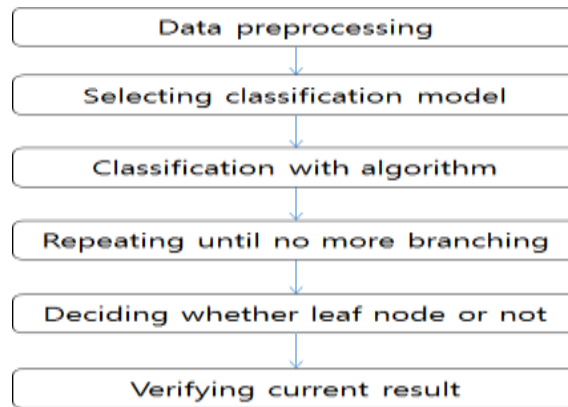


Figure 5. The Classification flow with decision tree algorithm

The first step calculates entropy when nodes are ramified and entropy before classification.

$$E(x) = \sum_{i=1}^n \psi(i) \div \log \psi(i) \tag{9}$$

$$E^{old} = (T(x) \times \log(T(x)) - E(x)) - \log(2) \tag{10}$$

The next step calculates new entropy after classification and then information profit quantity per branch case. The final step calculates information acquisition rate.

$$E^{old}(x) = \left(\frac{\sum_{j=1}^n (E(x_j) - T(x_j) \times \log(T(x_j)))}{\log(2)} \right) \tag{11}$$

4. IGAS for Analyzing the Gas Facility Status

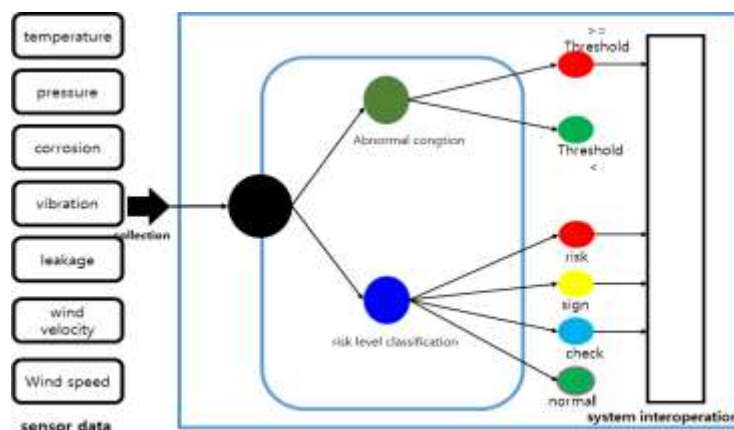


Figure 6. The Principle of IGAS

Our research has developed IGAS (Intelligent Gas Analysis System), which adopted our proposed intelligent method, for analyzing the status of LPG facility. IGAS adopted our intelligent method. IGAS collect remote seven kinds of sensor data that are

temperature, pressure, corrosion, vibration, gas leakage, wind velocity, and wind speed. Therefore, results of our IGAS will grow reliability using these collected sensor data

The IGAS would carry out parallel analysis process depending on whether emergency response or not such as Figure 6. If sensor data is outside the range of threshold, our system must execute abnormal cognition for processing emergency response such as urgency alarm. Otherwise, our system should decide autonomically risk level of facility. The result of risk level shows web-based system and Mobile system and notifies user as shown Figure 7 and Figure 8. This kind of system can be predictive maintenance by recommending safety decision



Figure 7. Web Interface for IGAS

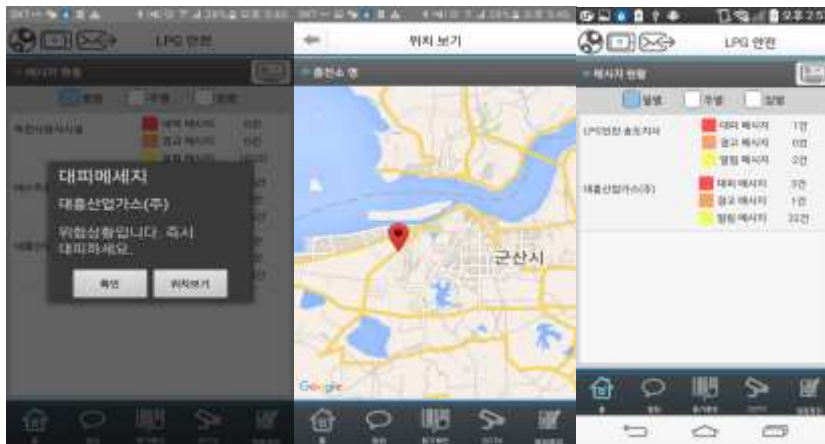


Figure 8. Mobile Interface for IGAS

5. Conclusions and Future Works

In gas accidents, human is damaged directly because most cause is careless handling and inadequate safety equipment, most type is explosion, fire and rupture. Therefore, an advanced safety process is required for enforcing safety management of LPG. Although correct progress direction in advances safety process is necessary for many devices, methods and system, it is to require analysis method that many different data are analyzed in parallel but each measuring data is analyzed in individual.

This paper preferentially aims to check the feasibility of machine learning analysis in order to apply a safety of gas facility, and devise method using artificial intelligent algorithm in order to manage total safety analysis that can analyze simultaneously many different data. At First, the feasibility study must be generally selected target gas facility, collected what kinds of risk factor, and then considered the appropriated machine learning method. We want to know corrosion status of pipeline using Bayesian algorithm about pipelines. In our feasibility test, inference results using Bayesian network show about 80% accuracy by validation of experts. According to cases, this Bayesian inference model is necessary to optimize algorithm for gas safety domain in order to improve accuracy. Furthermore, our study use decision model. Our decision tree model is about 92.2% accuracy by 10-cross validation method. In this paper, more than 90% accuracy would mean usable model in this domain for future classification.

Next, our research develops total risk analysis algorithm with a combination method between classification and clustering algorithm. Firstly, we applied and modified clustering EM algorithm, which is classified roughly into expectation procedure and maximization procedure, in order to generate efficiently risk class name. Secondly, this study customized decision tree algorithm, which create decision model through six procedures, for recommending autonomously facility status. This method is to mark the beginning of analysis method for detecting cause and increasing safety about LPG facilities.

Finally, we developed IGAS (Intelligent Gas Analysis System) for small gas facility version. The purpose of IGAS with our proposed algorithm must be complex algorithm whereby it is to minimize weakness and maximized advantages in gas safety environment. IGAS collect various kinds of sensor data and analysis the status of LPG facility. . If sensor data is outside the range of threshold, our system must execute abnormal cognition for processing emergency response. Otherwise, our system should decide autonomously risk level of facility. In future work, we will design and develop risk map to gas facility using IGAS.

Acknowledgments

These should be brief and placed at the end of the text before the references. This work was supported by Energy Efficiency & Resource Core Technology Program of Korea Institute of Energy and Planning (KETEP) granted financial resource from the Ministry of Trade, Industry & Energy, Republic of Korea (No. 20162220100150)

References

- [1] M.I. Moussa and T. Eid, "Risk Management for Chlorine Producing Factory in Egypt, Australian Journal of Basic and Applied Science, vol. 1, no. 2, (2007), pp. 239-248.
- [2] A. Daniel and J. F. Louvar, "Chemical Risk Analysis: Fundamentals with Applications, Prentice Hall, (1990).
- [3] M. Goeble and L. Gnuenwald, "A Survey of Data Mining and Knowledge Discovery Software Tools", ACM SIGKDD Explorations, vol. 1, no. 1, (1999), pp 20-33.
- [4] S. L. Tsang and S. Clarke, "Mining User Model for Effective Adaption of Context Aware Applications", International Journal of Security and Its Application, vol. 2, (2008), pp 5-12.
- [5] V. Henrique, K. M. Celso, R. G. Ivan, C. F. Tiago, and R. P. José, "Leak detection in petroleum pipelines using a fuzzy system", Journal of Petroleum Science and Engineering, vol. 49, no 3-4, (2005), pp 223-238.
- [6] C. Zangle and J. Nannerer, "Data Mining Applications in the Petroleum Industry, IBM Round Oak Publishing, (2003).
- [7] A. M. Akib, N. Saad, and V. Asirvadam, "Pressure Point Analysis for Early Detection System", Proceedings of IEEE 7th International Colloquium on Signal Processing and Its Applications, Penang, Malaysia, (2010) 103-107.
- [8] J. S. Oh, "An Analysis Intelligent Methodology for Finding Gas Leak Location", INFORMATION-An International Interdisciplinary Journal, vol. 16, no. 3(B), (2013), pp. 3297-3302.

- [9] J. S. Oh, "Safety Monitoring System Design for LPG Supply Facilities", The Asian International Journal of Life Sciences, no. 12, (2015), pp. 167-177.

Author



Jeong Seok Oh, He is a principal researcher at gas safety R&D institute in Korea Gas Safety Corporation. He received M.Sc degree and Ph.D degree both in computer science department from Soongsil University in 1998 and 2006 respectively. He has published more than 40 articles in convergence field between gas safety and IT, His main research interests include industrial IoT, risk analysis based on combination safety technology and information technology.

