

Intelligent Methods to Extract Knowledge from Process Data in the Industrial Applications

Young-Kwang Woo, Hyeon Bae, Sungshin Kim, and Kwang Bang Woo*

School of Electrical and Computer Engineering, Pusan National University, Busan, Korea

*Automation Technology Research Institute, Yonsei University, Seoul, Korea

Abstract

Data are an expression of the language or numerical values that show some features. And the information is extracted from data for the specific purposes. The knowledge is utilized as information to construct rules that recognize patterns or make a decision. Today, knowledge extraction and application of that are broadly accomplished for the easy comprehension and the performance improvement of systems in the several industrial fields. The knowledge extraction can be achieved by some steps that include the knowledge acquisition, expression, and implementation. Such extracted knowledge is drawn by rules with data mining techniques. Clustering (CL), input space partition (ISP), neuro-fuzzy (NF), neural network (NN), extension matrix (EM), etc. are employed for the knowledge expression based upon rules. In this paper, the various approaches of the knowledge extraction are surveyed and categorized by methodologies and applied industrial fields. Also, the trend and examples of each approaches are shown in the tables and graphs using the categories such as CL, ISP, NF, NN, EM, and so on.

Key words : Knowledge & Rule Extraction, Clustering, Neuro-Fuzzy, Extension Matrix

1. Introduction

The artificial intelligence is an algorithm that organizes rules from the expert's experienced knowledge and implement the results when it is necessary. Here a term of 'knowledge' is used or divided by a concept as information. The concept of knowledge is strictly defined by the data viewpoint. Information is usually collected from process data. Knowledge is accumulative know-how through systematic practical use of information [1]. Information becomes quarry of competitive power as itself [2]. If information is developed by knowledge, competitive power is more solidified.

This paper wished to compare and examine methods about studied knowledge extraction so far hereupon. We first grasped existent research tendency that was used for the knowledge abstraction on the whole and classified according to algorithm.

The application types are categorized by the big items including Clustering [3], Input space partition [4], Neuro-Fuzzy modeling [5], and Neural Network [6] etc. And the other techniques such as extension procession, high order division etc. are surveyed and necessary information of the techniques is collected. Through such methodic classification, we wished to examine kinds of intelligence style method that is used for knowledge abstraction on the whole. Also, the methods are classified corresponding to application fields with methodologies. If the suitable method for the goal is selected, good results in the system control or optimization should be guaranteed. To get the proper approach with respect to the application goal is very important for the best performance.

This paper surveys the knowledge extraction methods that

have been studied so far. We wish to provide the information for seeding the suitable data mining method according to the user's purposes.

2. Data Mining and Knowledge Extraction

2.1 Needs of Knowledge in Industrial Fields

The process of discovering knowledge from data is achieved through combining disciplines such as machine learning, statistic regression, soft computing, database theory, and visualization. The data mining research quoted, we need to keep in mind that neither of the various data mining methods is a panacea for solving real world problems involving hundreds of thousands of highly dimensional records.

Based on this quote one realizes the need to discover the features and characteristics of different knowledge discovery techniques, and the need to select the suitable ones for selected application. In order to do so, data mining supportive techniques should also be investigated, that include:

- Statistical Approaches
- Artificial Intelligent Approaches
- Soft Computing
- Knowledge Engineering

Development of an intermediate mechanism to bridge the different systems, considering the variations of the legacy databases and the variations of the data types adopted by different data mining techniques [7].

The scope of this research includes a study of knowledge extraction methods and how it may be best utilized using

knowledge extraction tools to improve day-to-day operations.

The objective of this research is to deliver an understanding of the impact of knowledge-based decision-making techniques on the manufacturing enterprise performance. This approach can be applied in the following four competencies [8], [9]:

Design: the ability to rapidly deploy and reconfigure manufacturing production capacity based on demand for goods. Rapidly changing consumer demands are forcing companies to accelerate and streamline the design and deployment of production processes. Rapid new product introductions in response to changing market demand is a critical competitive advantage and a key to growth. Having a design competency which ensures new products are not only launched on time, but that those launches are cost effectively executed is a key to translating that top-line growth into bottom-line results.

Operate: optimization of process yield and consistency throughout the enterprise. Plant productivity has always been a focus in manufacturing. As manufacturers consolidate through acquisitions and have new facilities to operate worldwide, they must further figure out how to get products made with consistent quality and efficiency at each manufacturing location. This is where initiatives like lean manufacturing drive out excess, achieving non-stop operations for maximum efficiency and throughput of production and where techniques like Six Sigma reduce variability in processes to ensure peak quality.

Maintain: efficient management of all company assets materials, processes, and employees to ensure non-stop operations, and optimum asset productivity. Without such a solid, efficient foundation, it is not possible to withstand the rigors of this fast paced environment, where growth and profits are demanded simultaneously.

Synchronize: tight coupling of a manufacturing operation into the greater supply chain, both upstream and downstream. This fourth competency is best achieved only after the other three competencies (Design, Operate, Maintain) are firmly in place. Only then is the plant truly ready to be fully coupled into an e-commerce driven supply chain.

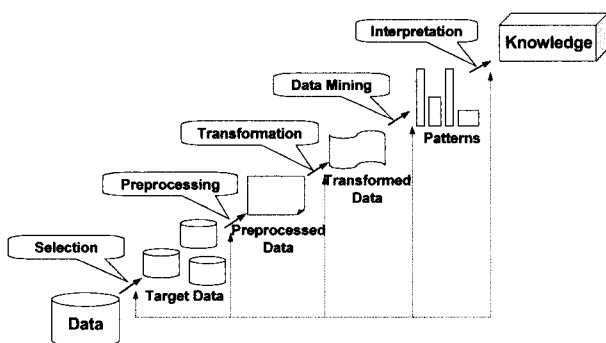


Fig. 1. Procedure of data mining and knowledge extraction.

2.2 Trend of Knowledge Extraction Techniques

The method of knowledge abstraction is one of the most intelligence style methods that have been developed so far. It selected 26 of 108 papers about knowledge abstraction which was inserted to IEEE about the present condition of various kinds knowledge abstraction methods for 10 years recently and analyzed tendency. In table 1, since NF was introduced at early 1990s, it is used important method to abstract knowledge and clustering (CL) [10] and input space partition method (ISP) are applied much. Also, method by extension procession [11], high order division [12], decision tree (DT) [13] etc. are studied unrelentingly in addition to these main methods. Significantly, input space partition method that was concept that just classify input space is reapplied changing by conditional input space partition that reflect expert's knowledge in late 1990s.

Table 1. Trend of paper publicatin of knowledge extraction

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
CL	1	1	2	3	1	3	6	3	1	0
ISP	5	2	4	6	6	2	2	2	3	5
NF	0	1	1	4	2	6	4	2	3	0
NN	2	0	1	3	0	0	2	0	1	0
etc.	1	2	2	2	0	0	0	2	3	0

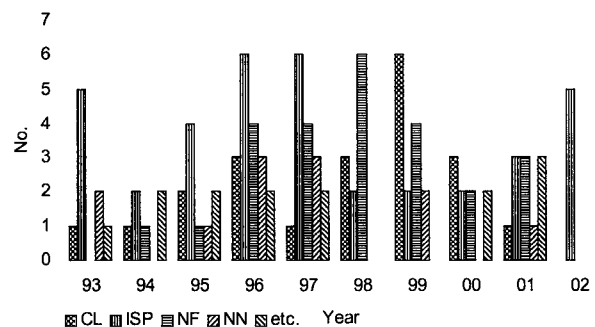


Fig. 2. Trend graph of knowledge extraction methods.

3. Comparison in Application Fields

In this chapter, We examine how knowledge abstraction methods were applied in awareness, diagnosis, environment, stocks system etc. We examine about methods used mainly by each area, and about reason the methods used

3.1 Application in Control Systems

Control systems used clustering or optimizing methods which make approximate rules and optimize these.

N. T. Gurman announced contents that compare 'Rule Net' model that add Clustering to information abstraction part with model suited Rule Net to fuzzy logic [3].

Z. M. Yeh developed 2 axis automatic control device that is linear and time-variable [14].

D. Kim *et al.* used Lamackin co-adapted fuzzy logic for the Backer-upper control of truck [15].

H. J. Cho *et al.* announced results that compared Genetic fine tuning method with Ziegler-Nichols method and manual tuning method [16].

A. Lekova *et al.* proposed method that optimize number of classified rule after finding point where individual input spaces are overlapped in each other according to input situation that happen in intersection by genetic algorithm [17].

B. J. Choi *et al.* announced method that reduce number of rule by genetic algorithm after making approximate rule using distance form input to control line in sliding mode control of pendulum [18].

F. Song *et al.* proposed method that derive rule from clustering after composing optimized control table in pendulum control [19].

O. F. Nelles *et al.* developed system that controls afterward state of pressure using current speed, fuel inflow, pressure [20].

M. Setnes derived optimized rule from changing number of clustering about each variable to control pressure while fermentation processes [21].

Table 2. Control system.

Field	Character	Tec.	Ref.
Robot navigation	Compose rules using fuzzy clustering and optimization with GA	CL	[3]
Contour control	2 axis automatic control device that is linear and time-variable	ISP	[14]
Control of truck	Compose rules using fuzzy clustering and optimization with GA	NF GA	[15]
PID control	Compose rules using fuzzy clustering and optimization with GA	NF GA	[16]
Crossroad	Compose rules using fuzzy clustering and optimization with GA that reduction number of rules	NF GA	[17]
Pendulum control	Reduce number of rule by genetic algorithm after making approximate rule using distance form input to control line	NF GA	[18]
	Compose rules using table of optimal control	CL	[19]
Pressure control	Optimize number of classified rule after finding point where individual input spaces are overlapped in each other according to input situation that happen in intersection by genetic algorithm.	NF GA	[20]
	Compose rules using fuzzy clustering	CL	[21]

3.2 Application in Recognition Systems

We used mainly division method of input space by condition that expert's knowledge is reflected in cognitive system.

K. Nozake *et al.* announced method that derive rules from heuristic method. He found rules by defining weight of

membership function as 'a multiplication' after dividing input space by temporary *k* piece and composing number of membership function as *k* [4].

S. H. Huang *et al.* proposed method that compose rules after sorting input by fuzzy input value and express by several binary this and learning neural net according to knowledge [22].

S. Chiu applied method that optimize membership function by Gradient Descent algorithm after derive initial rules from clustering of input space to cognitive system [23].

N. R. Pal *et al.* announced algorithm that reduce rules by confirm result after making initial rules and adjusting these by Gradient Descent algorithm [13].

Table 3. Recognition system.

Field	Character	Tec.	Ref.
Pattern recognition	Find rules by defining weight of membership function as 'a multiplication' after dividing input space by temporary <i>k</i> piece and composing number of membership function as <i>k</i>	ISP	[4]
	Compose rules after sorting input by fuzzy input value and express by several binary this and learning neural net according to knowledge	ISP NF	[22]
	Compose rules using fuzzy clustering	CL	[23]
	Compose rules with decision tree	DT	[13]

3.3 Application in Diagnostic Systems

Method used mainly to remove rules that less influenced in decision-making among model's already composed rules in diagnostic system.

T. Pfeufer *et al.* propose remove neuron that affect is few in decision-making by Neuro-fuzzy to grasp actuator's state that regulate the automotive speed directly [24].

M. Ayoubi *et al.* define transfer function using automotive tire partiality and characteristic telegram frequency and developed method to remove neuron that affect is few in decision-making by training for Neuro-fuzzy [25].

Table 4. Diagnosis system.

Field	Character	Tec.	Ref.
Actuator of automobile	Remove neurons that affect is few in decision-making by training for Neuro-fuzzy.	ISP NF	[24]
Tire pressure of automobile	Remove neurons that affect is few in decision-making by training for Neuro-fuzzy.	ISP NF	[25]

3.4 Application in Environmental Systems

Environment expert's knowledge is reflected to rule in environment system.

M. Cococcioni *et al.* proposed to optimize rule by genetic algorithm that decide early rule by Clustering to diagnose

pollution level of the sea using MERIS data [26].

E. F. Carrasco *et al.* developed the Fuzzy inference system that do fetters in expert's knowledge apply to oxidation state of wastewater processing [27].

J. P. Steyer *et al.* designed fuzzy controller that use expert's knowledge by input design for aversion process control of industrial waste water processing plant and through sensor values [28].

S. Dash *et al.* announced contents about chemical treatment process diagnosis. Sort ProFile that could get from sensor by pattern matching method and composed rule [29].

A. Muller *et al.* using fuzzy logic pattern in waste water process control that disturbance of wastewater treatment process under dangerous input condition [30].

Table 5. Apply for environments

Field	Character	Tec.	Ref.
Mesurement of sea polution	Compose rules using fuzzy clustering and optimization with GA	CL GA	[26]
Oxidation state of wastewater processing	Sort profiles that could get from sensor by pattern matching method and composed rule	ISP	[27]
Chemical treatment process diagnosis	Design fuzzy controller based expert's knoledge	ISP	[28]
Chemical process	Sort profiles that could get from sensor by pattern matching method and composed rule	ISP	[29]
Waster water trement process	Use fuzzy logic pattern in waste water process control that disturbance of wastewater treatment process under dangerous input condition.	ISP	[30]

3.5 Application in Stocks Prediction Systems

Several methods are used in stocks system.

D. D. Ettes developed system that do singleton fuzzy model and TS-fuzzy model by estimate model and optimize model by genetic algorithm for stocks. Used 20 price data in every day of Dutch AEX for simulation [31].

M. Setnes *et al.* announced research about analysis of stock market using TS -fuzzy model. Dutch AEX -other two model structure (Scenario model, Prediction model)-are used, and expressed linguistic rule using fuzzy clustering about modelling of stock index [32].

H. Kgaloozadeh *et al.* show result that long-term estimate of Iran stock market that appropriate by Neural Network. Performance compares with ARIMA model [33].

S. C. T. Chou *et al.* apply intelligence transaction system that forecast buying and selling signal using Neural Network. Used knowledge in form of reasoning to compose structure of Neural Network [34].

S. S. Lam developed new stock market timing system that graft genetic algorithm to fuzzy expert System. Genetic algorithm used in optimization of fuzzy buying and selling

rule. Suitable fuzzy rule is selected from information based on genetic algorithm [35].

Z. Pan *et al.* using Neuro-fuzzy system for banking modelling and estimation. Link of 2 - 3 layer of neural network are expresses rule through trained back propagation [36].

Table 6. Apply for stock market.

Field	Character	Tec.	Ref
Genetic Fuzzy Modeling	Compose rules using fuzzy clustering in input partition space	CL	[31]
Fuzzy Modeling	Compose rules using fuzzy clustering in input partition space	CL	[32]
Forecast Tehrna index	Each layer of neural network are expresses rule	NN	[33]
Decision making system in stock market	Each layer of neural network are expresses rule	NN	[34]
Timing system	Compose rules based upon expert's knoledge	ISP	[35]
Neuro-Fuzzy system	Link of 2nd-3rd layer of neural network are expresses rule	NF	[36]

4. Conclusion

Modern enterprise is first of all, vision and constant quest and continuous improvement, as a way of life, demands that an enterprise's practices, methods, and technologies constantly change. Also request about method that extract knowledge automatically is increasing rapidly. Methods for real time efficient management and optimization of system are studied continuously through that information circulation ability of computer connected with Information communication technology. Recognized about present condition of methods and the application field that is used in knowledge extraction field depending on these trend. conclusively, in case methods of input space partition and Clustering etc. much used in expert's knowledge base system, and result of system's condition is not detected certainly, method to optimize rule that using neural network and genetic algorithm used and sometimes two case was used together.

Reference

- [1] Jay Liebowitz, *Knowledge Management Handbook*. pp. 1-3, CRC Press, 1999.
- [2] Michael J. A. Berry, *Data Mining Techniques*. pp. 5-7, John Wiley & Sons, Inc. 1997.
- [3] Nadine Tschichold-Gurman, "The neural network model RuleNet and its application to mobile robot navigation," *Fuzzy Sets and Systems*, vol. 85, pp.287-303, 1997.
- [4] Ken Nozaki, Hisao Ishibuchi, and Hideo Tanaka, "A simple but powerful heuristic method for generating

- fuzzy rules from numerical data," *Fuzzy Sets and Systems*, vol. 86, pp. 251-270, 1997.
- [5] C. T. Lin, *Neural Fuzzy Control Systems with Structure and Parameter Learning*. World Scientific, pp. 8-10, 16-20, , 1994.
- [6] Sigeru Omatu, Marzuki Khalid, and Rubiyah Yusuf, *Neuro-Control and its Applications*. Springer, pp. 7-26, 1995.
- [7] J. Han, *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, 2001.
- [8] M. J. Shaw, *Information-Based Manufacturing with the Web*, *Int. J. Flexible Manuf. Systems*. Vol. 12, pp. 115-129. 2000.
- [9] R. Strauss, T. Hummel, *Information Technology, Business Process Re-engineering, and Lean Management in the Self-Organizing Fractal Company*, *Proc. IEEE Annual Int. Engr. Management Conf.*, pp. 287-292, 1995.
- [10] Yan Shi and Masaharu Mizumoto, "An improvement of neuro-fuzzy learning algorithm for tuning fuzzy rules," *Fuzzy Sets and Systems*, vol. 118, pp. 339-350, 2001.
- [11] X. Z. Wang, Y. D. Wang, X. F. Xu, W. D. Ling, and D. S. Yeung, "A new approach to fuzzy rule generation-fuzzy extension matrix," *Fuzzy Sets and Systems*, vol. 123, pp. 291-306, 2001
- [12] S. Abe and Ming-Shong Lan, "A Function Approximator Using Fuzzy Rules Extracted Directly From Numerical Data," *Proceedings of International Joint Conference on Neural Networks*, vol. 2, pp. 1887-1892, 1993.
- [13] N. R. Pal and S. Chakraborty, "Fuzzy rule extraction from ID3-type decision trees for real data," *IEEE Transactions on Systems, Man and Cybernetics, Part B*, vol. 31, no. 5, pp. 745-754, 2001.
- [14] Zong-Mu Yeh, "A cross-coupled bistage fuzzy controller for biaxis servomechanism control," *Fuzzy Sets and Systems*, vol. 97, pp. 265-275, 1998.
- [15] Daijin Kim and Han-Pyul Lee, "An optimal design of neuro-FLC by Lamarckian co-adaptation of learning and evolution," *Fuzzy Sets and Systems*, vol. 118, pp. 319-337, 2001.
- [16] Hyun-Jun Cho, Kwang-Bo Cho, and Bo-Hyeun Wang, "Fuzzy-PID hybrid control-Automatic rule generation using genetic algorithms," *Fuzzy Sets and Systems*, vol. 92, pp. 305-316, 1997.
- [17] A. Lekova, L. Mikhailov, D. Boyadjiev, and A. Nabout, "Redundant fuzzy rules exclusion by genetic algorithms," *Fuzzy Sets and Systems*, vol. 100, pp. 235-243, 1998.
- [18] Byung-Jae Choi, Seong-Woo Kwak, and Byung Kook Kim, "Design of a single-input fuzzy logic controller and its properties," *Fuzzy Sets and Systems*, vol. 106, pp. 299-308, 1999.
- [19] F. Song, S. M. Smith, and C. G. Rizk, "A general cell state space based TS type fuzzy logic controller automatic rule extraction and parameter optimization algorithm," *The 25th Annual Conference of the IEEE Industrial Electronics Society*, vol. 3, 1999.
- [20] O. Nelles, M. Fischer, and B. Muller, "Fuzzy rule extraction by a genetic algorithm and constrained nonlinear optimization of membership functions," *Proceedings of the Fifth IEEE International Conference on Fuzzy Systems*, vol. 1, pp. 213-219, 1996.
- [21] M. Setnes, "Supervised fuzzy clustering for rule extraction," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 4, pp. 416-424, 2000.
- [22] Samuel H. Huang and Hao Xing, "Extract intelligible and concise fuzzy rules from neural networks," *Fuzzy Sets and Systems*, vol. 132, pp. 233-243, 2002.
- [23] S. Chiu, "Method and software for extracting fuzzy classification rules by subtractive clustering," *Biennial Conference of the North American Fuzzy Information Processing Society*, pp. 461-465, 1996.
- [24] Thomas Pfeufer and Mihiar Ayoubi, "Application of a hybrid neuro-fuzzy system to the fault diagnosis of an automotive electromechanical actuator," *Fuzzy Sets and Systems*, vol. 89, pp. 351-360, 1997.
- [25] Mihiar Ayoubi and Rolf Isermann, "Neuro-fuzzy systems for diagnosis," *Fuzzy Sets and Systems*, vol. 89, pp. 289-307, 1997.
- [26] M. Cococcioni, G. Corsini, M. Diani, R. Grasso, B. Lazzarini, and F. Marcelloni, "Automatic extraction of fuzzy rules from MERIS data to identify sea water optically active constituent concentration," *Annual Meeting of the North American Fuzzy Information Processing Society*, pp. 546-551, 2002.
- [27] E. F. Carrasco, J. Rodriguez, A. Punal, E. Roca, and J. M. Lema, "Diagnosis of acidification states in an anaerobic wastewater treatment plant using a fuzzy-based expert system," *Control Engineering Practice*, Available online 28, 2003.
- [28] J.-P. Steyer, M. Estaben, and M. Polit, "Fuzzy control of an anaerobic digestion process for the treatment of an industrial wastewater," *Proceedings of the Sixth IEEE International Conference on Fuzzy Systems*, vol. 3, pp. 1245-1249, 1997.
- [29] Sourabh Dash, Raghunathan Rengaswamy, and Venkat Venkatasubramanian, "Fuzzy-logic based trend classification for fault diagnosis of chemical processes," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 347-362, 2003.
- [30] A. Muller, S. Marsili-Libelli, A. Aivasidis, T. Lloyd, S. Kroner, and C. Wandrey, "Fuzzy control of disturbances in a wastewater treatment process," *Water Research*, vol. 31, no. 12, pp. 3157-3167, 1997.
- [31] Drs. Dennis Ettes, "Trading the stock markets using Genetic Fuzzy Modeling," *CIFER*, pp. 22-25, 2000.
- [32] M. Setnes and O. J. H. van Drempt, "Fuzzy Modeling in Stock-Market Analysis," *Proceedings of the IEEE/IAFE 1999 Conference on Computational Intelligence for Financial Engineering*, pp. 250-258, 1999.
- [33] Hamid Kgaloozadeh and A'li Khaki Sedigh, "Long Term Prediction of Tehran Price Index (TEPIX) Using Neural Networks," *Joint 9th IFSA World Congress and 20th NAFIPS International conference*, vol. 1, pp. 25-28,

2001.

- [34] Seng-Cho Timothy Chou, Chau-Chen Yang, Chi-Huang Chen, and Feipei Lay, "A Rule-Based Neural Stock Trading Decision Support System," *Proceedings of the IEEE/IAFE 1996 Conference on Computational Intelligence for Financial Engineering*, pp. 148-154, 1996.
- [35] S. S. Lam, "A Genetic Fuzzy Expert System for Stock Market Timing," *Proceedings of the 2001 Congress on Evolutionary Computation*, vol. 2, pp. 410-417, 2001
- [36] Zouhong Pan, Xiadi Liu, and Olugbenga Mejabi, "A Neural - Fuzzy System for Forecasting," *Proceedings of the Thirtieth Hawaii International Conference on System Sciences*, vol. 5, pp. 549-558, 1997.
-

Young-Kwang Woo

He is currently pursuing the M.S degree at Pusan National University. His research interests include fuzzy logic and manufacturing systems.



Hyeon Bae

He received the M.S. degree in electrical engineering from Pusan National University in 2001. He is currently pursuing the Ph.D. degree at Pusan National University. His research interests include intelligent modeling and data mining.



Sungshin Kim

He received the B.S. and M.S. degrees from Yonsei University, and the Ph.D. degree in electrical engineering from Georgia Institute of Technology, Atlanta, in 1984, 1986, and 1996, respectively.

He is currently an Assistant Professor in the School of Electrical & Computer Engineering, Pusan National University. His research interests include intelligent control, fuzzy logic control, manufacturing systems, and data mining.

Phone : +82-51-510-2374

Fax : +82-51-513-0212

E-mail : sskim@pusan.ac.kr



Kwang Bang Woo

He received the B.S. and M.S. degrees from Yonsei University, and the Ph.D. degree in electrical engineering from Oregon State University, in 1957, 1959, and 1964, respectively.

He is currently an Research Professor in Automation Technology Research Institute, Yonsei University. His research interests include intelligent control, system and factory automation.