A General Situation Assessment Model Based on Rule Extraction

Xiao-dan Zhang
School of Information Science & Engineering
Northeastern University
Shenyang, 110004
P.R.China
zhangsheny@neuera.com

Hai Zhao
School of Information Science & Engineering
Northeastern University
Shenyang, 110004
P.R.China
zhangsheny@neuera.com

Zhan-sheng Duan
School of Information Science & Engineering
Xi'an Jiaotong University
Xi'an, 710049
P.R.China
zhangsheny@neuera.com

Abstract - *In this paper, a general situation assessment model based on knowledge obtained is proposed, that is, gathering multi-source data, obtaining knowledge automatically using modified Rough Set algorithm, and the general situation assessment being realized by utilizing the modified blackboard model in dynamic circumstance. This model is applied into bus protection of hydroelectrical simulation system to solve high complexity, redundancy of rule, and lower generality of former system. Through the experiment in Jilin Fengman hydroelectrical simulation system, validity and application value of the method has been proved.*

Keywords: Situation assessment; blackboard model; rough set; hydroelectric simulation; bus protection

1 Introduction

Information fusion began in 1970's, and developed rapidly in 1980's. Recently, it has been improved rapidly not only in theory, but also in application. Situation assessment, which can utilize different sensors to offer redundant or complementary information, is high layer function model of the information fusion system. It is used in some research areas, such as industry process watch, military affairs direct, avigation traffic control, and medical treatment, etc. [1-4]。

There are many situation assessment methods, the hot topic is the situation assessment method based on rule. The rule extraction is the kernel. In the paper, a new and general situation assessment method based on rule extraction is proposed, that is, obtaining real-time data by multi-sensors, getting attribute rule and decision rule by modified rough set, and realizing situation assessment by general and effective blackboard model.

At last, the model is applied in bus protection of Jilin Fengman hydroelectrical simulation system successfully.

2 General Model of Situation Assessment Based on Rule Extraction

Situation assessment is an automatic information disposal process which can get information via multi-sensors from the environments, and makes decision and evaluation of the current situation or its possible development trend. The process mainly includes situation abstract and assessment.

It is a challenging work to extract right and simple knowledge from incomplete, fuzzy and ever-changing information. In this paper, a general fusion model is proposed which is based on modified rough set and blackboard model to solve the above problems. It is showed as Fig.1.

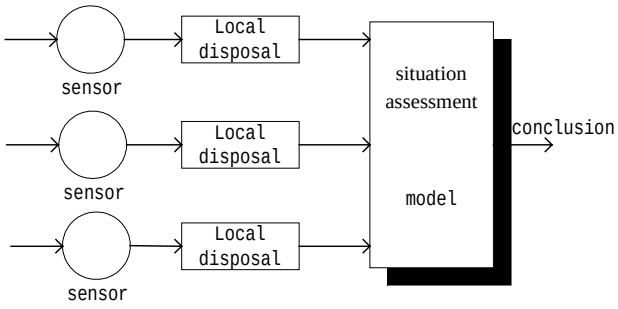

Fig.1. The whole fusion model

In the model, the information is acquired by multi-sensors, then pretreated (including filtering, denoising, compression, etc.), at last sent to the modified blackboard model for knowledge extraction, inference, and exporting evaluation result. Meanwhile, heuristic algorithms for attribute reduction and decision rule reduction are proposed to reach the aim to acquire knowledge and simplification rules.

3 Blackboard Model of Situation Assessment

It is difficult to get accurate and reliable information during the progress of information extraction in usual information extraction. So the system needs judge by analyzing technology parameter, data information, and various environments information. Blackboard model is a very effective situation assessment method[5]. People did much improvement in different application fields, such as real-time blackboard model and distributed blackboard model for solving real-time problems[6][7], repository model for solving generality and extensibility based on blackboard. To solve generality, a new and general blackboard model is proposed in the paper, it is showed as Fig.2.

In this model, the blackboard is divided into three layers: parameter level, attribute level, situation level. The information and various middle suppose brought during the progress of evaluation are recorded in the relevant blackboard level, and information can be shared by the knowledge source of the same and the higher layer.

According to different knowledge levels used in the system and relevant information layer, the knowledge source of the model is divided into parameter level, attribute level, and situation level. These knowledge sources include all relevant field knowledge that is needed in the evaluation progress. The knowledge sources of the model get information from the structure of the blackboard on the same layer, the reference results modify the blackboard information, and afford information to solve problems. In the course, the communication between blackboard model and knowledge source is accomplished through simulation model and response model. When the information on the blackboard has chance, the stimulator and dispelled confliction are accomplished by stimulator model. The mid-conclusion or the final conclusion are accomplished by the respond model.

In the model, reference engine accomplishes the work, for instance, the choice of reference mechanism, the explanation mechanism which brings out explanation. The function of general control model can coordinate the reference engine, respond model, simulation model, corresponding model and the management of the knowledge source.

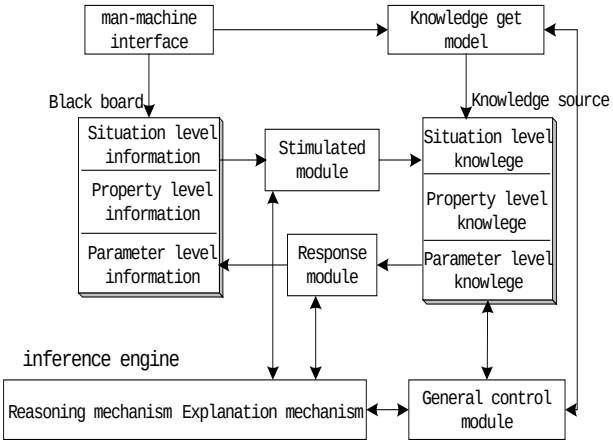

Fig.2. General situation assessment blackboard model

4 Progress of Rule Acquired Based on Improved Rough Set

Usually, the rule is offered by field experts, but this method has many disadvantages, for example, it is only applied in some fields, lack of universality and generality, and the independence is not well. If the data brought by the system to acquire a series of rule knowledge possessing generality, and the main idea of the rule extraction is utilizing equivalence class, attribute reduction and decision reduction to reach the objective of excavation rule and simplification rule. This paper advances the decision rule reduction arithmetic.

4.1 Rule Acquired Structure Based on Rough Set

Rough set theory is proposed by Poland scientist and professor Pawlak initially in 1982, it is mainly used in expressing, studying, and inducting uncertainty data and knowledge [8-9]. Rough set has been applied in many research areas widely, such as machine study, knowledge discover, decision sustainment. And it can analyze and dispose inexact , unwholesome data, and find the hidden relation. In this paper, an improved rough set method is applied in analysis mass data to acquire field knowledge automatically in the progress of evaluation. The progress is showed as Fig.3.

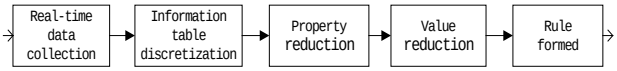

Fig.3 The progress of knowledge acquired based on rough set

In Fig.3., real-time collection model accomplishes data collection, and expresses various data that produced during the system running.

Definition1: knowledge express system is orderly pairs $S=<U,C,D,V,f>$, meanwhile, U is nonempty and finite set, named corpora, which is gathering data or example set in different time. C is condition attribute set, D is decision attribute set, $R=C\cup D$ is named as attribute set. V is attribute value set, which appoints every object attribute value of U . $f:U\times R\rightarrow V$ is an information function, which appoints every object: x 'attribute value of U .

Because the disposed data by rough set method are discrete, and the gathered data usually are continuous, the results of gathering data need discret. In the model of information table discretion, the ways of the usual discretion method are supervising and non-supervising diverse. Usually non-supervising discrete includes two methods: constant width interval method and equi-frequency interval method. Though non-supervising diverse method is very simple, its effection is not well because the relation, between condition attribute and decision attribute, is ignored. Supervising diverse method is often used in practical application, including ChiMerge arithmetic, statistical arithmetic, and filter progress, etc.

In the discretion information table, the condition attribute and the acquired data in different time are redundancy possibly, and the basic point is attribute redundancy and derivative reduction concept, that is judging whether every equivalence class essential, and how to detract unnecessary relation. Reduction is composed of attribute relation and decision relation. During the progress of attribute relation, discernible matrix advanced by A.Skowron is used to express knowledge and reduct the information table.

The main steps are showed as followed:

Step1, calculating discernible matrix M(S) of the information table;

Step2, calculating discernible function fM(S) corresponding to discernible matrix M(S);

Step3, calculating the minimal disjunctive normal form of discernible function fM(S), among them every extraction distribution is corresponding to a reduction.

In the rule formed model, attribute rule is advanced from original data according to the above steps, and rule set acquired at every level through heuristic arithmetic of decision rule.

During the evaluation progress, the system data are judged and the corresponding evaluation results, according to decision rule, already exist.

4.2 Utilizing the Heuristic Arithmetic of Decision Rule to Extract Minimization Rule Arithmetic

Rough set describes uncertainty relation of knowledge based on uncertainty relation of knowledge express system. In the knowledge system, if example x_i and record x_j are equal to the value then x_i and x_j are considered equal based on attribute set, the set of all equivalence record based on attribute set A is defined as equivalence class. The record belongs to a class attributes to the uniform equivalence class, and the class becomes class partition that A is based on attribute set R, shows as $Ulind(R)=(R_1,R_2,...,R_n)$, $[x]_R$ shows equivalence class that $Ulind(R)$ includes X.

When $X \gamma U$, and X is the union, and X is named that A is defined, and is A' accurate set.，or X is named that A can't be defined，also is named the rough set of A.

In knowledge system, every example represents a decision rule. But the value of every condition attribute in decision rules is not necessary. The reduction of decision rule is to cancel unnecessary condition attribute value of decision rules, that is, calculating the kernel and simplify every rule. Usually, after attribute reduction, the redundancy value of rule is eliminated.

In $S=(U,C,D,V,f)$, defines $Des(C_i)$ the value of equivalence class C_i to the equivalence relation of condition attribute C, $Ulind(C)=\{C_1,C_2,L,C_n\}$.

The extraction of decision rule is to acquire the rule of the express form as followed:

$Des(a_1)\gamma Des(a_2)\gamma L\gamma Des(a_n)\gamma Des(d)$, meanwhile, $Des(a)$ is a certain equivalence class value of condition attribute a, $Des(a)$ is a certain equivalence class value of decision attribute.

After attribute reduction and the repeat rules union, the decision rule is simplified.

In the knowledge system, if condition attribute is $C=\{a_1,a_2,L,a_n\}$, decision attribute is d, the value of attribute a_1 is $\{a_{1,1},a_{1,2},L,a_{1,n_1}\}$, Then the equivalence class of a_1 is $ind(a_1)=\{E_{1,1},E_{1,2},L,E_{1,n_1}\}$, $Des(E_{1,1})=a_{1,1}$, $Des(E_{1,2})=a_{1,2}$, ..., $Des(E_{1,n_1})=a_{1,n_1}$.

Every decision rule of decision table simplification includes the kernel extraction and the rule simplification.

$Des(a_{i,j})$ expresses the value a_i of rule j, $|Des(a_{i,j})|$ expresses the equivalence class of attribute a_i, according to $Des(a_{i,j})$,the related equivalence class is found indefinable matrix, showed as:

$Des(a_{1,j})\mathrm{I}Des(a_{2,j})\mathrm{I}L\mathrm{I}Des(a_{n,j})rDes(d_j)$,

Definition2: if $Des(a_{1,j})\mathrm{I}Des(a_{2,j})\mathrm{I}L\mathrm{I}Des(a_{n,j})rDes(d_j)$ then the value of attribute a_1 can be omitted in rule. $Des(a_{2,i})\mathrm{I}L\mathrm{I}Des(a_{n,i})$ is a reduction rule, noted as $red(i)$. The kernel of rule i is: $core(i)=\mathrm{I}red(i)$.the number of reduction rule is $k(i)$.

The minimal solution decision rule is detracted, and acquired to reduction rules. The arithmetic is heuristic and solving minimal simplification arithmetic。

In the irrecognizable matrix defined, the element in $E_{i,j}$ is classed according to the equivalence class $Ulind(d)$ of decision attribute, for example:

$E_{1,1}=\{x_1,x_2,x_3\}$, $d(x_1)=d(x_2)=2$, $d(x_3)=1$, the arithmetic is descried as followed:

1) Extracting the element in $E_{i,j}$ which only has a class based on the uniform record $Ulind(d)$, such as $E_{2,1}=\{\{x_1,x_2,2\}\}$, $E_{2,2}=\{\{x_1,x_2,1\},\{x_3,1\}$ has the uniform record in two types. The uniform record based on $Des(d)$ is union operation, after that, the related attribute value is one dimension agent, and can acquire one value reduction rule.

2) Dispose the item which includes many uniform record $V(d)$.

$E_{i_1,j},E_{i_2,j}$ only have one uniform record element, and union uniform record. At last, binary value simplification rule is acquired, If $E_{i_1,j},E_{i_2,j}$ are multi-record elements, and the example x of the element is not included in the former rules, then ternary simplification.

3) The value simplification rules, until all examples x are included in the rules. simplification method is like (2);

4) Gathering all the rules, and the minimal decision table is formed.

5 Experiment Analysis

5.1 Situation Assessment System of Bus Protect

Fengman hydroelectical simulation system is developed jointly by Neu and Neuera Company. In the tide calculation model of the simulation system, current is in steady state. Moreover the math model of diagnose is concluded according to the changeable situation of current and voltage when diagnose of bus happens. Multi-diagnose running models include the calculation of short circuit, open-circuit current, etc. [10].

5.2 Bus Data Collection

In the progress of bus data real-time collection, model LH, identify LB, current-carrying capacity LEI, voltage-carrying capacity LEU、power-carrying capacity LEP are all static state value, but A,B,C, triphase current LAI, LBI, LCI, and triphase voltage LAU、LBU、LCU and work station D on the bus are dynamic values. But bus accident and fault are only related to current and voltage directly.

In Table 1 U is the observation node, the unit of current is A，the unit of voltage is KV. Attribute D is the attribute of current circuit work station, it is showed as followed:

$$D=\begin{cases}0 & normal\\ 1 & single\ phase\ breakage\ groupshort\ circuit\\ 2 & inter\ phase\ groupshort\ circuit\\ 3 & inter\ phase\ shortcircu\ it\\ 4 & triphase\ shortcircu\ it\\ 5 & single\ phase\ breakage\\ 6 & inter\ phase\ breakage\\ 7 & triphase\ breakage\end{cases} \tag{1}$$

Table 1. Original collection information table of bus system

U	LA	LB	LCI	LAU	LB	LC	D
P1	301	299	298	221	220	219	0
P2	1	1	1	0	0	0	0
P3	300	257	257	220	80	80	1
P4	257	257	300	90	90	220	1
P5	257	300	257	90	220	90	1
P6	300	300	128	210	220	55	2
P7	300	129	300	220	55	210	2
P8	300	200	300	200	55	210	2
P9	300	128	128	220	65	65	3
P10	129	129	300	70	70	220	3
P11	129	300	129	70	220	70	3
P12	423	465	418	74	70	73	4
P13	0	310	330	225	226	231	5
P14	300	300	0	230	224	230	5
P15	300	0	300	220	220	220	5
P16	0	300	0	220	220	220	6
P17	300	0	0	220	220	220	6
P18	0	0	0	220	220	220	7

5.2 Bus Information Table Discrete

During the process of disperse the bus information table, this paper adopts the method that Nguyen and Skowron extracts minimal number set with heuristic arithmetic, and chooses disperse save point is that: current disperse （20，500，2250），voltage disperse （150）. So disperse rule of attribute LAI、LBI、LCI can be acquired:

$$a,b,c=\begin{cases}0 & I<20A\\ 1 & 20\le I<500A\\ 2 & 500\le I<2250A\\ 3 & I\ge 2250A\end{cases} \tag{2}$$

The disperse rules of attribute LAU、LBU、LCU are:

$$d,e,f=\begin{cases}0 & U<150kV\\ 1 & U\ge 150kV\end{cases} \tag{3}$$

5.4 Bus Information Table Reduction

After the former information table disperses, the information table can be detracted, which includes three steps: condition attribute reduction, value reduction, and minimal rules reduction. According to Table 1, the condition attribute reduction is involved mostly during the progress of reduction. In this example, a modified condition attribute reduction method is adopted. The decision table after reduction is showed as Table 2:

Table 2.The information table after reduction

U	LAI	LBI	LCI	LAU	D
P1	1	1	1	×	0
P2	0	0	0	0	0
P3	1	3	3	×	1
P4	3	3	1	×	1
P5	3	1	3	×	1
P6	1	1	2	×	2
P7	1	2	1	×	2
P8	2	1	1	×	2
P9	1	2	2	×	3
P10	2	2	1	×	3
P11	2	1	2	×	3
P12	3	3	3	×	4
P13	0	1	1	×	5
P14	1	1	0	×	5
P15	1	0	1	×	5
P16	0	1	0	×	6
P17	1	0	0	×	6
P18	0	0	0	1	7

5.5 Bus Protect Rule Producing

Correlative and related bus protect rules are formed on the base of Table 2;

Natural situation decision rules: abc triphase current is between 20A and 500A; abc triphase current is less than 20A and the phase of a or b or c is less than 150V;

Inter phase short circuit decision rules: a phase current is between 20A and 500A ,and bc phase current is bigger than 2250A;b phase current is between 20A and 500A, and ac phase current bigger than 2250A;c phase current is between 20A and 500A, and ab phase current is bigger than 2250A;

Single phase short decision rules: a phase current is between 500A and 2250A, and bc phase current is between 20A and 500A; b phase current is between 500A and 2250A, and ac phase current is between 20A and 500A; c phase current is between 500A and 2250A, and ba phase current is between 20A and 500A;

Inter phase short circuit rules: a phase current is between 20A and 500A, and bc phase current is between 500A and 2250A; c phase current is between 20A and 500A, and ba phase current is between 500A and 2250A;

Triphase short circuit decision rules: abc triphase current is all bigger than 2250A;

Single phase breakage decision rules: a phase current is less than 20A, and bc phase current is between 20A and 500A; b phase current is less than 20A, and ac phase

current is between 20A and 500A; c phase current is less than 20A, and ba phase current is between 20A and 500A; Inter phase breakage decision rules: a phase current is between 20A and 500A, and bc phase current is less than 20A; b phase current is between 20A and 500A, and ac phase is less than 20A; c phase current is between 20A and 500A, and ab phase current is less than 20A;

Triphase breakage decision rules: abc triphase current are all less than 20A, and phase voltage of a, b, c is bigger than 150V.

We can find from the above rules that the number of the condition attribute included by all rules is under 4, in fact, only the number of two rule' condition attribute is 4, the other only have three condition attribute. The former repository needs judge more than 8 pieces of condition, thus the complementary and the redundancy knowledge are reduced.

5.6 Bus Situation Assessment Progress

The repository used in blackboard model is acquired through knowledge acquiring progress based on rough set theory. So the current and factual work station(such as nomination, short circuit, disconnection, etc.) are concluded through inference according to current data, and a real emulation situation is provided through starting accident protect reaction set.

The repository built in the evaluation progress is the bus protect rules in section 5.4. In the progress of inference, the paper adopts direct inference method, which is decided by rough set' character. The condition attribute is combined in the combinatorial theory of rough set, and through combinatorial operation 'and', 'or', and 'not', to conclude decision rules. The repository is extracted successfully from database, the final decision is concluded utilizing decision rules which are formed by combination.

6 Experiment Contrast with Traditional Evaluation Method

In the factual application of the simulation system, the fault can be set and removed automatically through teacher beach according to preconcerted strategy under formal work station. in the experiment analysis, the two sub system, traditional hydroelectrical fault diagnose model (sensor-threshold method) and the situation assessment model is formed first for the contrast of experiment data, and two client pc run at the same time to contrast capability effectively under the condition of the uniform set. Secondly, the example of union net work station is chosen , and various faults had been randomly and automatically set and removed. Table 3, 4 is the part of the experiment results.

Table 3 Experiment result using sensor-threshold

Work station	class	example	Judge right	Judge wrong	Right probability
Active	O	54	45	9	83.3%
	F_1	128	63	65	49.2%
	F_2	119	59	60	49.6%
	F_3	79	17	62	21.5%
	F_4	95	50	45	52.6%
	F_5	210	101	109	48.1%
Active	O	232	197	35	84.9%
	F_1	356	99	257	38.5%
	F_2	516	121	395	23.4%
	F_3	424	215	211	50.7%
	F_4	206	75	131	36.4%
	F_5	456	169	287	37.1%

Table4 Experiment result using the situation assessment

Work station	species	example	Right judgment	Wrong judgment	Right probability
Active load 50%	O	50	47	3	94.0%
	F_1	198	136	62	68.7%
	F_2	135	110	25	81.5%
	F_3	85	76	9	89.4%
	F_4	96	83	13	86.5%
	F_5	125	99	26	79.2%
Active load 98%	O	270	246	24	91.1%
	F_1	451	375	76	83.1%
	F_2	525	485	40	92.4%
	F_3	366	287	79	78.4%
	F_4	295	256	39	86.8%
	F_5	321	280	41	87.2%

7 Conclusion

In this paper, a general situation assessment model based on knowledge obtained is proposed, that is, gathering multi-source data, obtaining knowledge automatically using modified Rough Set algorithm, and the general situation assessment being realized by utilizing the modified blackboard model in dynamic circumstance. Based on rough set theory, the method acquires bus protect rules from vast real-time data, and solves the design and realization about the high redundancy of rule and low generality of the former bus system. In the evaluation progress, the situation evaluation has been adopted that based on the modified blackboard model, bus protect rules has been utilized according to improved heuristic producing arithmetic. Meanwhile. Through the evaluation and decision-made of the bus situation , the independence of individual model and generalized degree have been improved in the simulation system, the design of the system has been simplified. The advantage of the method has been proved in the bus protect application of Fengman hydroelectrical simulation system.

References

[1] Hall D. Mathematical Techniques in Multi-sensor Data Fusion [M]. Norwood: Artech House, 1992. 235–238.

[2] Waltz E, Llinas J. Multi-sensor data fusion [M]. Norwood: Artech House, 1991. 101-105.
[3] Llinas, J. Blackboard concepts for data fusion applications [J]. International Journal of Pattern Recognition and AI, 1993,7(2): 285-308.
[4] Johnson, J, et al. Real time blackboard for sensor fusion [C]. Proceedings of SPIE the Inter. Society for Optical Engineering Sensor Fusion, Washington, 1989,1100: 61-71.
[5] Harmon, S. Y, et al. Sensor data fusion through a distributed blackboard [C]. Proceedings of 1986 IEEE inter. Conf. On Robotics and Automation, Sanfrancisco, 1986: 1449-1454.
[6] Shahbazian, E, et al. Fusion of imaging and non-imaging data for surveillance aircraft [C]. Proceedings of SPIE-the Inter. Society for Optical Engineering, Washington, 1989,1100: 61-71.
[7] Shahbazian, E, et al. Multi-agent data fusion workstation(MADFW) architecture [C]. Proc of SPIE- The Inter. Society for Optical Engineering, Sensor Fusion: Architecture, Algorithm and Applications II, Washington, 1998, 3376: 60-68.
[8] Pawlak Z. Rough sets[J]. International Journal of Computer and Information Sciences,1982 ,11 :341 – 356
[9] Pawlak Z. Rough sets: Theoretical Aspects of Reasoning about Data[A]. Boston:Kluwer Academic Publishers, 1991, 1-10
[10] Xu L Y, Du Q D. Application of neural fusion to accident forecast in hydropower station. Proceedings of the Second International Conference on Information Fusion. Vol2 Sunnyvale, 1999, pp. 1166-1171.