

THE APPLICATION OF SPECIALIZED DATA STRUCTURES FOR SCADA DIAGNOSTICS

Abstract. The general characteristics and working methods relating to the specialized data structure template, such as m -tuples based on the ordered set of arbitrary power are considered in the paper. This structure has been applied during the diagnostic model research for the diagnosis search area on the classification space of the Expert System knowledge base. The method for determining the functional dependency between element index of ordered set that is subset of base set power and m -tuple of base set indexes is examined in detail. The example of the software screenshot designed to the work with m -tuples is given. The advantages of working with this specialized structure to increase the efficiency of intellectual operations for SCADA diagnostics are shown.

Keywords: diagnostics, fault tolerance, SCADA, data organization structure, m -tuples.

Problem statement. At this time, SCADA (Supervisory Control And Data Acquisition) systems are widely spread due to the intensification of production, implementation of the high technologies and the increasing complexity of the technological process and equipment [1, 2]. The distinctive feature of this system class is a supervisory control based on HMI (Human Machine Interface). The system reliability is not only affected by hardware-software safety and effectiveness of the control algorithms of TCO (Technological Control Object), but also the validity of actions performed by operators to elaborate the system diagnostic messages. This inappropriate diagnostic subsystem work leads to the time deficit, emergency situations and unscheduled downtimes.

According to high requirements to ensure the safety of production, there is an actual task to improve the methods of SCADA diagnostics in real time. One of the ways to solve this problem is the development of the independent outside automated “expert” based on the methodology of Expert Systems. It allows for a reduction in the flow of diag-

nostic messages directed to operator and for providing comprehensive, preprocessed diagnostic information. In that regard, the accuracy and effectiveness of operator actions of SCADA work are increased. Also, the reliability and fault tolerance of system are raised.

Analysis of publications on the topic research

Analysis of the diagnosis subsystems of well-known SCADA systems such as WinCC (Siemens, Germany), SPPA-T3000 (Siemens, Germany) and Genesis 32 (Iconics, USA) [3, 4, 5] allows us to draw an important conclusion. This is that as a rule, automated maintenance of SCADA systems can be performed by using “hot” backups of systemically important units, with automatic switching to reserve equipment in case of failure.

User interface proposed by the diagnostic subsystems of the most common SCADA displays vast amounts of diagnostic information to monitor the states and analyze errors. However, the low-level system diagnostics are usually presented similar to the service that is provided by utilities and operating system functions and developmental environments.

It is easier to analyze diagnostic information about some possible reasons of failure by using built-in system diagnostic mechanisms that trace errors and provide memory dumps. However, even by using software debuggers it is difficult to track periodically occurring errors due to the mistiming of interaction between two or more SCADA real time processes in the wide area network.

At this time, the Expert System methodology is not widely used in diagnostic subsystems of the most common SCADA.

At the same time, many developing toolkits for Expert Systems, such as Expert System shells, specialized environments and programming languages are proposed to professionals [6].

Among the best known Expert System shells, there are such as G2 (Gensym, USA), TDC Expert (Honeywell, USA), RT Expert (Integrated Systems Inc., USA), ILOG Rules (ILOG, France) [7].

The distinctive feature of the Expert System methodology usage for SCADA diagnostic subsystems is the application of both algorithmic methods and heuristic methods, the possibility of application methods that allows forming accurate logical conclusion under conditions of uncertainty [8].

Formulation of aims of the article

The purpose of the article is development of specialized data structure template such as m -tuples based on the ordered set of arbitrary power. This structure is applied to form the diagnosis search area on the classification space of the Expert System knowledge base for SCADA failure diagnostics. The task statement and diagnostic model using Expert System methodology are described in detail by the author in the paper [9].

Main part of the article

Suppose we have a finite base set X which is in ascending order:

$$X = \{x_1, \dots, x_i, x_{i+1}, \dots, x_n\} = \{x_i \mid 1 \leq i \leq n, \forall i \in [1, n) \rightarrow x_i < x_{i+1}\},$$

where i – sequential number of element x of the set X ; $n = |X|$ – number of elements or power of set X ; $<$ – comparison operator of the set X elements that might be defined depending on the element type of the set X ; 2^X – set of all subsets of X ; $|2^X| = 2^n$ – power of set of all subsets.

$$2^X = \{\emptyset, Y_1^n, \dots, Y_m^n, \dots, Y_n^n\} = \left\{ Y_m^n \mid 1 \leq m \leq n, k_m^n = |Y_m^n| = \frac{n!}{(n-m)! m!} \right\},$$

where Y_m^n – subset of the set of all subsets, the elements of which are m -tuples containing the elements of set X . These elements are sorted by using a right-handed search of their indexes in m -tuples, beginning from the lower boundary of index value, changing to the upper one; k_m^n – power of the subset Y_m^n .

$$Y_m^n = \left\{ y_{m,j}^n \mid y_{m,j}^n = (x_{i_1}, \dots, x_{i_\eta}, \dots, x_{i_m}), 1 \leq \eta \leq m, \alpha_{m,\eta}^n \leq i_\eta \leq \beta_{m,\eta}^n, \alpha_{m,\eta}^n = \eta, \right. \\ \left. \beta_{m,\eta}^n = n - m + \eta, \forall \eta \in [1, m) \rightarrow x_{i_\eta} < x_{i_{\eta+1}} \right\},$$

where $y_{m,j}^n$ – m -tuple; It is an element of set Y_m^n that contains the elements of base set X ; m – tuple length; j – tuple index (sequential number) in set of tuples Y_m^n ; η – element index within the tuple $y_{m,j}^n$; i_η – tuple element index $y_{m,j}^n$ within the base set X ; $\alpha_{m,\eta}^n$ – lower boundary of changing i_η in m -tuple $y_{m,j}^n$; $\beta_{m,\eta}^n$ – upper boundary of changing i_η in m -tuple $y_{m,j}^n$.

Suppose we have a set of indexes I which is in ascending order of elements of set X ; J_m^n – a set of element indexes of set Y_m^n which is in ascending order; I_m^n – a set of index tuples of base elements which correspond to tuples of set Y_m^n ; $i_{m,j}^n$ – a tuple of indexes of base elements in the tuple $y_{m,j}^n$.

$$I = \{1, 2, \dots, n\} = \{i \mid 1 \leq i \leq n\},$$

$$J_m^n = \{1, 2, \dots, k_m^n\} = \{j \mid 1 \leq j \leq k_m^n\},$$

$$I_m^n = \{i_{m,1}^n, i_{m,2}^n, \dots, i_{m,k_m^n}^n\} = \{i_{m,j}^n \mid i_{m,j}^n = (i_1, \dots, i_\eta, \dots, i_m), 1 \leq j \leq k_m^n, i_\eta \in I,$$

$$1 \leq \eta \leq m, \eta \leq i_\eta \leq n - m + \eta, \forall \eta \in [1, m) \rightarrow i_{\eta+1} - i_\eta \geq 1\}.$$

Consider the example of expansion of the set of all subsets 2^X into the sets Y_m^n for $n=7$ (see Tab. 1).

Table 1

Representation the set of all subsets 2^X by using sets Y_m^n for $n=7$

m	Y_m^n	k_m^n
1	$Y_1^7 = \{a, b, c, d, e, f, g\}$	7
2	$Y_2^7 = \{(a, b), (a, c), (a, d), (a, e), (a, f), (a, g), (b, c), (b, d), (b, e), (b, f), (b, g), (c, d), (c, e), (c, f), (c, g), (d, e), (d, f), (d, g), (e, f), (e, g), (f, g)\}$	21
3	$Y_3^7 = \{(a, b, c), (a, b, d), (a, b, e), (a, b, f), (a, b, g), (a, c, d), (a, c, e), (a, c, f), (a, c, g), (a, d, e), (a, d, f), (a, d, g), (a, e, f), (a, e, g), (a, f, g), (b, c, d), (b, c, e), (b, c, f), (b, c, g), (b, d, e), (b, d, f), (b, d, g), (b, e, f), (b, e, g), (b, f, g), (c, d, e), (c, d, f), (c, d, g), (c, e, f), (c, e, g), (c, f, g), (d, e, f), (d, e, g), (d, f, g), (e, f, g)\}$	35
4	$Y_4^7 = \{(a, b, c, d), (a, b, c, e), (a, b, c, f), (a, b, c, g), (a, b, d, e), (a, b, d, f), (a, b, d, g), (a, b, e, f), (a, b, e, g), (a, b, f, g), (a, c, d, e), (a, c, d, f), (a, c, d, g), (a, c, e, f), (a, c, e, g), (a, c, f, g), (a, d, e, f), (a, d, e, g), (a, d, f, g), (a, e, f, g), (b, c, d, e), (b, c, d, f), (b, c, d, g), (b, c, e, f), (b, c, e, g), (b, c, f, g), (b, d, e, f), (b, d, e, g), (b, d, f, g), (b, e, f, g), (c, d, e, f), (c, d, e, g), (c, d, f, g), (c, e, f, g), (d, e, f, g)\}$	35
5	$Y_5^7 = \{(a, b, c, d, e), (a, b, c, d, f), (a, b, c, d, g), (a, b, c, e, f), (a, b, c, e, g), (a, b, c, f, g), (a, b, d, e, f), (a, b, d, e, g), (a, b, d, f, g), (a, b, e, f, g), (a, c, d, e, f), (a, c, d, e, g), (a, c, d, f, g), (a, c, e, f, g), (a, d, e, f, g), (a, d, e, g), (a, d, f, g), (a, e, f, g)\}$	21

	$(b, c, d, e, f), (b, c, d, e, g), (b, c, d, f, g), (b, c, e, f, g), (b, d, e, f, g), (c, d, e, f, g)\}$	
6	$Y_6^7 = \{(a, b, c, d, e, f), (a, b, c, d, e, g), (a, b, c, d, f, g), (a, b, c, e, f, g), (a, b, d, e, f, g), (a, c, d, e, f, g), (b, c, d, e, f, g)\}$	7
7	$Y_7^7 = \{(a, b, c, d, e, f, g)\}$	1

Consider the example of expansion of the set Y_4^7 into the elements (see Tab. 2).

Table 2

Expansion of the set Y_4^7 into the elements

J_4^7	Y_4^7	$I_{m,j}^n$			
		i_1	i_2	i_3	i_4
1	(a, b, c, d)	1	2	3	4
2	(a, b, c, e)	1	2	3	5
3	(a, b, c, f)	1	2	3	6
4	(a, b, c, g)	1	2	3	7
5	(a, b, d, e)	1	2	4	5
6	(a, b, d, f)	1	2	4	6
7	(a, b, d, g)	1	2	4	7
8	(a, b, e, f)	1	2	5	6
9	(a, b, e, g)	1	2	5	7
10	(a, b, f, g)	1	2	6	7
11	(a, c, d, e)	1	3	4	5
12	(a, c, d, f)	1	3	4	6
13	(a, c, d, g)	1	3	4	7
14	(a, c, e, f)	1	3	5	6
15	(a, c, e, g)	1	3	5	7
16	(a, c, f, g)	1	3	6	7
17	(a, d, e, f)	1	4	5	6
18	(a, d, e, g)	1	4	5	7
19	(a, d, f, g)	1	4	6	7
20	(a, e, f, g)	1	5	6	7
21	(b, c, d, e)	2	3	4	5
22	(b, c, d, f)	2	3	4	6

23	(b, c, d, g)	2	3	4	7
24	(b, c, e, f)	2	3	5	6
25	(b, c, e, g)	2	3	5	7
26	(b, c, f, g)	2	3	6	7
27	(b, d, e, f)	2	4	5	6
28	(b, d, e, g)	2	4	5	7
29	(b, d, f, g)	2	4	6	7
30	(b, e, f, g)	2	5	6	7
31	(c, d, e, f)	3	4	5	6
32	(c, d, e, g)	3	4	5	7
33	(c, d, f, g)	3	4	6	7
34	(c, e, f, g)	3	5	6	7
35	(d, e, f, g)	4	5	6	7

Method for determining the dependency $j \rightarrow i_{m,j}^n$ that was developed by the author is proposed to consider in this paper.

We have: $X = \{x_i \mid n = |X|, 1 \leq i \leq n, \forall i \in [1, n) \rightarrow x_i < x_{i+1}\}, m, j$.

We have to find: $i_{m,j}^n = (i_1^j, \dots, i_\eta^j, \dots, i_m^j)$.

Decision:

1. Let's perform the validity of the input data: X , m and j .

$$1 \leq m \leq n,$$

$$k_m^n = \frac{n!}{(n-m)! \cdot m!},$$

$$1 \leq j \leq k_m^n, \forall i \in [1, n) \rightarrow x_i < x_{i+1}.$$

2. Let's form Tab. 3 with dimension $(m+1) \times 4$ for calculation the parameters $\alpha_{m,\eta}^n$, $\beta_{m,\eta}^n$, $\mu_{m,\eta}^n$:

$$\forall \eta \in [1, m] \rightarrow (\alpha_{m,\eta}^n = \eta, \beta_{m,\eta}^n = n - m + \eta, \mu_{m,\eta}^n = m - \eta).$$

Table 3

Variable scopes for indexes i_η

η	$\alpha_{m,\eta}^n$	$\beta_{m,\eta}^n$	$\mu_{m,\eta}^n$	} m
1				
...	η	$n - m + \eta$	$m - \eta$	
m				

3. Let's form Tab. 4 with dimension $(\beta_{m,1}^n - \alpha_{m,1}^n + 4) \times (2m + 1)$ for calculation the parameters $C_{n-i_\eta}^{m-\eta}$, that define the count of elements of the set Y_m^n . These elements were deducted by using ordered a right-handed search of their indexes i_η in the scope:

$$(\forall \eta \in [1, m]) \ \& \ (\forall i_\eta \in [\alpha_{m,\eta}^n, \beta_{m,\eta}^n]) \rightarrow C_{n-i_\eta}^{m-\eta} = \frac{(n - i_\eta)!}{((n - i_\eta) - (m - \eta))! \times (m - \eta)!}$$

Table 4

Table for calculation the parameters $C_{n-i_\eta}^{m-\eta}$

η	1		2		m		1
$\mu_{m,\eta}^n$	m - 1		m - 2		0		2
i_η	i_1	$C_{n-i_1}^{m-\eta}$	i_2	$C_{n-i_2}^{m-\eta}$	i_m	$C_{n-i_m}^{m-\eta}$	3
$\alpha_{m,\eta}^n$	$\alpha_{m,1}^n$	$C_{n-\alpha_{m,1}^n}^{m-1}$	$\alpha_{m,2}^n$	$C_{n-\alpha_{m,2}^n}^{m-2}$	$\alpha_{m,m}^n$	1	} $\beta_{m,1}^n - \alpha_{m,1}^n$
...	
$\beta_{m,\eta}^n$	$\beta_{m,1}^n$	$C_{n-\beta_{m,1}^n}^{m-1}$	$\beta_{m,2}^n$	$C_{n-\beta_{m,2}^n}^{m-2}$	$\beta_{m,m}^n$	1	

4. Let's form Tab. 5 with dimension $(\beta_{m,1}^n - \alpha_{m,1}^n + 2) \times 2m$ for calculation the parameters $J_{m,\eta}^n(i_\eta)$:

Table 5

Table for calculation the parameters $J_{m,\eta}^n(i_\eta)$

i_1	$J_{m,1}^n(i_1)$	i_2	$J_{m,2}^n(i_2)$	i_m	$J_{m,m}^n(i_m)$	1
$\alpha_{m,1}^n$		$\alpha_{m,2}^n$		$\alpha_{m,m}^n$		} $\beta_{m,1}^n - \alpha_{m,1}^n + 1$
...			
$\beta_{m,1}^n$		$\beta_{m,2}^n$		$\beta_{m,m}^n$		

Calculation $J_{m,\eta}^n(i_\eta)$ for filling the Tab. 5 is performed by using the formulas (1), (2), (3).

5. Calculation i_η^j is performed by using the formulas (2), (5).

6. The rules for completing the algorithm are given in the formulas (6), (7).

The following formulas show the sequence of calculation performance:

$$J_{m,1}^n(i_1) := \sum_{\alpha_{m,1}^n}^{i_1} C_{n-i_1}^{m-1} = \begin{cases} J_{m,1}^n(i_1) := C_{n-\alpha_{m,1}^n}^{m-1}, & i_1 = \alpha_{m,1}^n \\ J_{m,1}^n(i_1) := J_{m,1}^n(i_1 - 1) + C_{n-i_1}^{m-1}, & i_1 \in (\alpha_{m,1}^n, \beta_{m,1}^n] \end{cases}, \quad (1)$$

$$i^j \leftarrow \min_{i_1=\alpha_{m,1}^n}^{\beta_{m,1}^n} (J_{m,1}^n(i_1)) \geq j, \quad (2)$$

$$(\eta \in (1, m]) \rightarrow ((i_{\eta-1}^j > \alpha_{m,\eta-1}^n) ? (i_\eta := i_{\eta-1}^j, j_{\min} := J_{m,\eta}^n(i_\eta) := \\ := J_{m,\eta-1}^n(i_{\eta-1}^j - 1) : (j_{\min} := 0)), \quad (3)$$

$$(i_\eta \in (i_{\eta-1}^j, \beta_{m,\eta}^n]) \rightarrow \\ J_{m,\eta}^n(i_\eta) := j_{\min} + \sum_{i_{\eta-1}^j+1}^{i_\eta} C_{n-i_\eta}^{m-\eta} = \begin{cases} J_{m,\eta}^n(i_\eta) := j_{\min} + C_{n-i_\eta}^{m-\eta}, & i_\eta = i_{\eta-1}^j + 1 \\ J_{m,\eta}^n(i_\eta) := J_{m,\eta}^n(i_\eta - 1) + C_{n-i_\eta}^{m-\eta}, & i_\eta \in (i_{\eta-1}^j, \beta_{m,\eta}^n] \end{cases} \quad (4)$$

,

$$i_\eta^j \leftarrow \min_{i_\eta=i_{\eta-1}^j+1}^{\beta_{m,\eta}^n} (J_{m,\eta}^n(i_\eta)) \geq j). \quad (5)$$

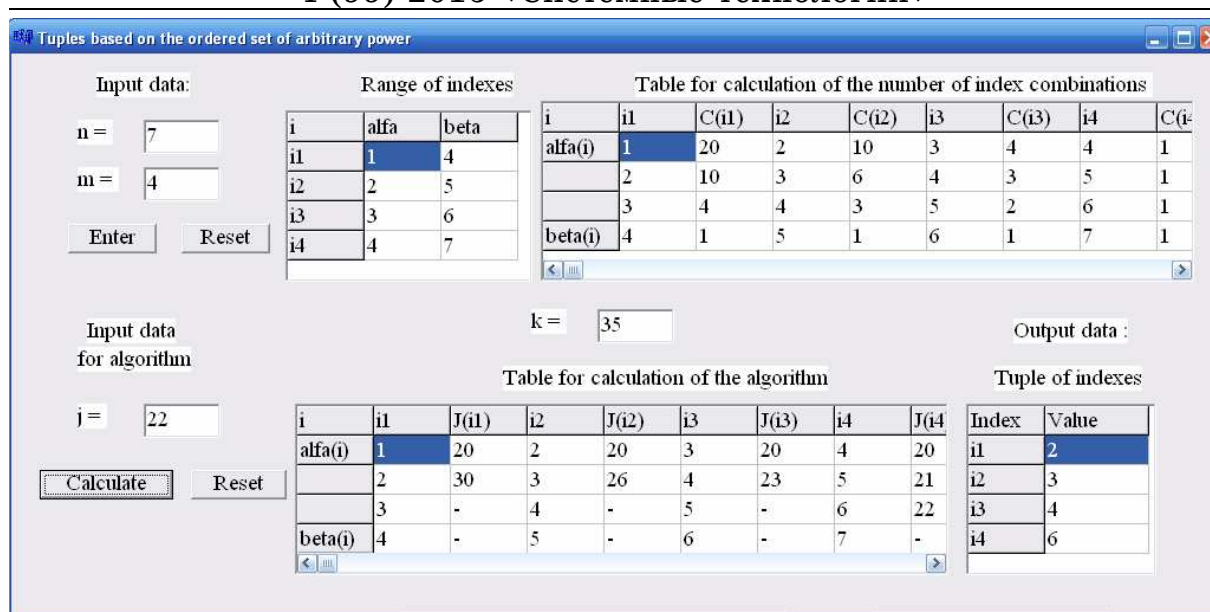
Final rules for the algorithm:

$$i_{\eta^*}^j \leftarrow (J_{m,\eta^*}^n(i_{\eta^*}) = j), \quad (6)$$

$$(\eta^* < m) \rightarrow (i_{\eta^*+h}^j := \beta_{m,\eta^*+h}^n, \quad h \in [1, m - \eta^*]) \quad (7)$$

Answer: $i_{m,j}^n = \{i_\eta^j \mid \eta \in [1, m]\}$.

The described method of determining dependency $j \rightarrow i_{m,j}^n$ for m -tuples based on the ordered set of arbitrary power was realized by using the software that is shown below (see Fig.1).

Figure 1 - Software screenshot for working with m -tuples

Conclusions and recommendations for further research

The described method of working with m -tuples based on the ordered set of arbitrary power allows us to significantly reduce the system resources spent for storing and processing the big data arrays. It allows for a generation in any m -tuple by using some consequence of formal rules. It should also be noted that the input data of the diagnostic subsystem of SCADA system are complex data structures. The proposed method minimizes the data processing duration due to the proposed indexation that allows us to work with the sequence of natural numbers instead of complex data structures.

The application of proposed specialized data structure during the formation diagnosis search area on the classification space of Expert System knowledge base allows us to increase the performance of intellectual operations, reduce the time of executing the procedures of conflict resolutions. Also, it allows for easily reconfiguration in the formal knowledge structures that may significantly reduce the time spent on the determining diagnosis about SCADA technical state.

As future-oriented studies in the given direction can be recommended for the development of strategies for working with the Expert System knowledge base with using data organization structure that was considered above for SCADA diagnostics.

REFERENCES

1. Stouffer K. Guide to supervisory control and data acquisition (SCADA) and industrial control systems security. // Recommendations of the National Institute of Standards and Technology. NIST Special publication 800-82 / Eds.: K. Stouffer, J. Falco, and K. Kent. — Gaithersburg, 2006. — 164 p.
2. Kudlak B.I. Purposes and base functions of SCADA, (2015), available at: <http://votum.ua/old/uk/publications/scada.htm> (accessed March 12, 2015).
3. Siemens. User manual SPPA-T3000. Diagnostic system / Siemens AG. — Erlangen, 2009. — 251 p.
4. The official site of SIMATIC WinCC SCADA System, (2015), available at: <http://automation.siemens.com> (accessed March 12, 2015).
5. The official site of Genesis32 SCADA system, (2015), available at: <http://iconics.com> (accessed March 12, 2015).
6. Giarratano J. Expert Systems: principles and programming/ J. Giarratano, G. Riley. — Moscow: Publishing house “Williams”, 2007. — 1152 p. (in Russian).
7. Popov E.V. Real time Expert Systems / E.V. Popov // Open Systems. 1995. — No. 2. available at: <http://www.osp.ru/os/1995/02/178608/> (accessed March 12, 2015, in Russian).
8. Jackson P. Introduction to Expert Systems / P. Jackson. — Moscow: Publishing house “Williams”, 2001. — 624 p. (in Russian).
9. Syrotkina O.I. Formation of the classification space of the Expert System knowledge base for SCADA failure diagnostics // Power Engineering, Control and Information Technologies in Geotechnical Systems. — 2015. — P. 179-184.