UDC 681.518.5

M. Alekseyev, I. Udovyk, O. Syrotkina
State Higher Educational Institution "National Mining University", Ukraine
19, Dmytra Yavornytshogo av., Dnipro, 49027

APPLICATION OF PREDICATE LOGIC FOR FAILURE DETECTION IN SCADA SYSTEMS

М. Алексєєв, І. Удовик, О. Сироткіна
Державний вищий навчальний заклад «Національний гірничий університет», Україна
пр. Дмитра Яворницького 19, Дніпро, 49027

ЗАСТОСУВАННЯ ЛОГІКИ ПРЕДИКАТІВ ДЛЯ ВИЯВЛЕННЯ ВІДМОВ У SCADA СИСТЕМАХ

We consider the task of failure detection and localization. It is based on the analysis of the information flow state change in the system. We suggest a structural and logical model to describe SCADA of any topology. It is possible to form diagnostic features of independent failure detection. They are based on the characteristic functions of three-valued logic. We determine the predicate system of knowledge representation to implement the method of SCADA diagnostics in the event of incomplete data.

Keywords: predicate system of knowledge representation, structural and logical model, three-valued logic, independent failure.

Розглядається задача виявлення та локалізації відмов у SCADA на основі аналізу зміни стану інформаційних потоків у системі. Пропонується структурно-логічна модель опису SCADA будь-якої топології. На основі характеристичних функцій тризначної логіки формуються діагностичні ознаки виявлення незалежної відмови. Визначається предикатна система подання знань для реалізації методу діагностики працездатності SCADA в умовах неповних даних / недостовірних даних.

Ключові слова: предикатна система подання знань, структурно-логічна модель, тризначна логіка, незалежна відмова.

Introduction

Considering the application of expert systems to diagnose SCADA performance, it should be noted that the relevant task is the development of a reliable and fast decision support system which significantly depends on the chosen method of knowledge representation [1–3].

All knowledge representation systems can be divided into the following main classes: declarative, procedural and special. Predicative systems refer to declarative knowledge representation systems. It is possible to distinguish procedures to find solutions (known as a generation mechanism) and procedures to optimize this search (management mechanism) for declarative knowledge representation systems.

Declarative systems are characterized by the universality of knowledge representation. The control mechanism, which determines the semantics of the declarative system and heuristic efficiency to search the solution, reduces the universality of knowledge representation. Thus, there is a contradiction between universality and efficiency of knowledge representation for declarative systems [1–3].

Publication analysis regarding topic research

We analyzed the latest research in the field of SCADA diagnostics using expert system methodology. It showed that today's expert diagnostic systems are focused on Technological Control Object (TCO) diagnostics. At the same time, they do not diagnose the whole SCADA system. Vast, intensive flows of low-level diagnostic information generated by SCADA causes significant difficulties in its processing by operational

personnel. Therefore, there is a need to implement expert systems as decision support systems for SCADA diagnostics in real time.

Problem statement

The pressing problem is automatic high-level SCADA diagnostics based on the methodology of expert systems in real time.

The aim of the research is to increase the quality of SCADA functioning by developing a method of automatic failure detection and localization in real time. It is based on the analysis of information flow change when passing through SCADA structural elements and hierarchy levels. To do this it is necessary to develop a knowledge representation system which can universally describe the following elements: SCADA structure of any topology; distribution of diagnostic features for independent failure detection through structural elements of different hierarchy levels; effective diagnosis search in real time.

Main part

Consider an example of a given fragment of SCADA structure (see Fig. 1) [4].

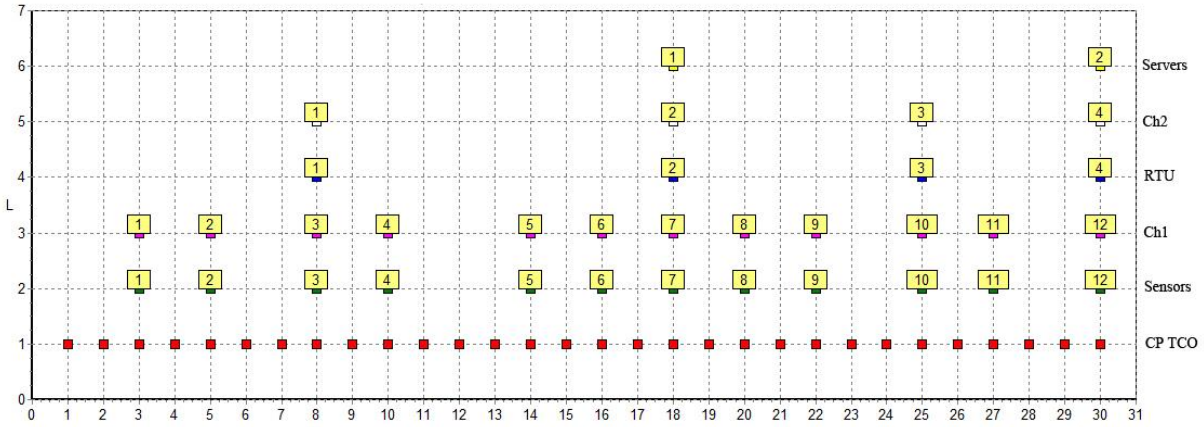

Fig. 1. SCADA structural scheme

The set of controlled parameters (CP TCO) at a point of time *t* is as follows:

$$X(t) = \{x_1(t), x_2(t), \ldots, x_i(t), \ldots x_{n(X)}(t)\}. \quad (1)$$

Controlled parameters are measured by sensors and are transmitted to RTUs through data transmission channels (Ch1). These controlled parameters are then transmitted to servers through Ch2.

We can apply the following non-decreasing sequences of natural numbers to describe the distribution of controlled parameters through structural elements of different hierarchy levels [5].

The sequence K_x determines the distribution of controlled parameters through sensors and Ch1:

$$K_x = k_1, k_2, \ldots, k_\mu, \ldots, k_{m_N}. \quad (2)$$

The sequence I_x determines the distribution of controlled parameters through RTUs and Ch2:

$$I_x = i_1, i_2, \ldots, i_j, \ldots, i_N. \quad (3)$$

The sequence M_k determines the distribution of sensors and Ch1 through RTUs and Ch2:

$$M_k = m_1, m_2, \ldots, m_j, \ldots, m_N. \quad (4)$$

We define certain predicates of connection between structural elements of different hierarchy levels using formulas (2) – (4).

The predicate of connection between CP TCO x_i and Sensor$_\mu$:

$$H_1(i,\mu) := ((i \le k_{m_N}) \,\&\, (\mu \le m_N))\,?((\mu > 1)\,?((i > k_{\mu-1}) \,\&\, (i \le k_\mu)) : (i \le k_\mu)) : 0, \quad (5)$$

where $m_j \in K_x$.

The predicate of connection between CP TCO x_i and RTU$_j$:

$$H_2(i,j) := ((i \le i_N) \,\&\, (j \le N))\,?((j > 1)\,?((i > i_{j-1}) \,\&\, (i \le i_j)) : (i \le i_j)) : 0, \quad (6)$$

where $i_j \in I_x$.

The predicate of connection between Sensor$_i$ and RTU$_j$:

$$H_3(i,j) := ((i \le m_N) \,\&\, (j \le N))\,?((j > 1)\,?((i > m_{j-1}) \,\&\, (i \le m_j)) : (i \le m_j)) : 0, \quad (7)$$

where $m_j \in M_k$.

We developed a method of automatic failure detection and localization in SCADA. The input data in this method are diagnostic matrix $D(t)$. The matrix is represented as a dump containing the diagnostic features of information flows. It is formed with the sample rate of data from sensors. The number of rows in this matrix corresponds to the number of SCADA hierarchy levels. The number of columns corresponds to the number of controlled parameters.

$$\begin{cases} D(t) = [d_{iL,iC}(t)] \\ d_{il,iC}(t) \in E_3, E_3 = \{0,1,2\} \\ iL = l(S_1) + 1 - l, \;\; 1 \le l \le l(S_1) \\ 1 \le iC \le n(X) \end{cases}, \quad (8)$$

where iL – the index of the matrix row $D(t)$ which corresponds to SCADA hierarchy levels l; iC – the index of the matrix column $D(t)$ which corresponds to the index of the controlled parameter $x_{iC}(t)$; $l(S_1)$ – the hierarchy level of servers; $n(X)$ – the number of controlled parameters.

The controlled parameter can have one of three states at each SCADA hierarchy level: “Absent,” “Non-reliable,” “Reliable.” These states can be described by using Post’s three-valued logic.

We apply the elementary function of three-valued logic φ_e – the characteristic function of the first kind with value e to analyze diagnostic matrix $D(t)$.

$$\phi_e(x) = \begin{cases} 1, & x = e, \;\; e \in E_3, \;\; E_3 = \{0,1,2\} \\ 0, & x \ne e, \;\; e \in E_3, \;\; E_3 = \{0,1,2\} \end{cases}. \quad (9)$$

We define diagnostic features for failure detection as follows:

a) A sufficient diagnostic feature of failure absence for the SCADA structural element (iL, iC) at a point of time t is:

$$\varphi_2(d_{iL,iC}(t)) = 1; \quad (10)$$

b) A necessary but insufficient diagnostic feature of failure detection for the SCADA structural element (iL, iC) at a point of time t is:

$$\neg\varphi_2(d_{iL,iC}(t)) = 1; \quad (11)$$

c) A necessary but insufficient diagnostic feature of failure detection due to the absence of controlled parameters at a hierarchy level (this corresponds to backbone nodes) or due to absence of data transmission process (this corresponds to data transmission channels Ch1/Ch2) for the SCADA structural element (iL, iC) at a point of time t is:

$$\varphi_0(d_{iL,iC}(t)) = 1; \quad (12)$$

d) A necessary but insufficient diagnostic feature of failure detection due to the unreliability of controlled parameters at a hierarchy level (this corresponds to backbone nodes) or due to the unreliability of data transmission (this corresponds to data transmission channels Ch1/Ch2) for the SCADA structural element (*iL, iC*) at a point of time *t* is:

$$\varphi_1(d_{iL,iC}(t)) = 1. \quad (13)$$

Analyzing diagnostic matrix $D(t)$ we can assert that no failures have been detected at a point of time *t* if the following expression is true for the first row (*iL* = 1) of diagnostic matrix $D(t)$ which corresponds to the server's hierarchy level $l(S_1)$:

$$\&_{iC=1}^{i_N} \varphi_2(d_{1,iC}(t)) = 1. \quad (14)$$

In general, the function of failure detection based on the analysis of diagnostic matrix $D(t)$ is as follows:

$$g_2(iL,\alpha,\beta,t) = \neg(\&_{iC=\alpha}^{\beta} \varphi_2(d_{iL,iC}(t))), \quad (15)$$

where *iL* – the index of the matrix row $D(t)$ which corresponds to SCADA hierarchy levels *l*; α, β – the initial and final ordinal numbers of controlled parameters which pass through the system's structural elements for the given hierarchy level *l*.

Consider the predicate $S(i,y,l)$ to form the criteria of diagnostic feature distribution through independent failures taking into account the characteristic attributes for each SCADA hierarchy level. This predicate determines the state *y* for the controlled parameter x_i at the hierarchy level *l*:

$$S(i,y,l) := (y = d_{l(S1)+1-l,i}). \quad (16)$$

Then the diagnostic feature of failure detection can be described by the predicate $S(i,y,l)$ as follows:

a) $S(i,2,l)$; b) $\neg S(i,2,l)$; c) $S(i,0,l)$; d) $S(i,1,l)$.

It should be noted that both diagnostic features of failure detection $\neg\varphi_2(d_{iL,iC}(t))$ and the function of failure detection in SCADA $g_2(iL,\alpha,\beta,t)$ do not distinguish independent and secondary failures. We assume that all the failures are independent at the lowest level l_{min} for the given controlled parameter when passing through SCADA hierarchy levels. Thus, all diagnostic features of failure detection refer to these features of **independent** failure detection at level *l* ($l < l_{min}$). We also assume that diagnostic features at hierarchy levels which correspond to data transmission channels Ch1/Ch2 are diagnostic features of **independent** failures.

Therefore, at this stage of diagnostic matrix $D(t)$ analysis we can assert the following:

- The absence of diagnostic features for failure detection at a certain SCADA hierarchy level is a sufficient condition that no failures have been detected at this hierarchy level;
- The presence of diagnostic features at hierarchy level l_{min} is a sufficient condition that there are independent failures at hierarchy level l_{min} and all the diagnostic features of failure detection refer to independent failures;
- The number of independent failures at a hierarchy level of sensors ($l_{min} = 2$) is equal to the number of diagnostic features for failure detection;

- In order to define the number of failures at a certain low hierarchy level ($l_{min} > 2$), the additional analysis of diagnostic matrix $D(t)$ is necessary because various diagnostic features can refer to the same failure;
- The presence of diagnostic features for failure detection at hierarchy level l ($l > l_{min}$) for $l \in L_1$ is a necessary but insufficient condition of having an independent failure from low hierarchy levels. It is necessary to have additional diagnostic criteria to consider it an independent or secondary failure;
- The absence of diagnostic features for independent failure detection at hierarchy level l ($l > l_{min}$) is a sufficient condition of having no independent failures at this hierarchy level;
- To define the number of independent failures when having diagnostic features for independent failure detection at hierarchy level l ($l > l_{min}$), it is necessary to conduct an additional analysis of diagnostic matrix $D(t)$ because different diagnostic features can refer to the same failure.

We define the lowest level of SCADA for failure detection l_{min} in accordance with SCADA structure (see Fig. 1).

$$\begin{aligned} iL_{\min} := {} & (g_2(1,1,i_N,t))? \\ & ((g_2(2,1,i_N,t))? \\ & \quad ((g_2(3,1,i_N,t))? \\ & ((g_2(4,1,i_N,t))? \\ & ((g_2(5,1,i_N,t))?5:4):3):2):1):0. \end{aligned} \tag{17}$$

If $iL_{min} = 0$, then no failures have been detected at a point of time t. Otherwise, the lowest hierarchy level of failure detection in SCADA is as follows:

$$l_{\min} = l(S_1) + 1 - iL_{\min}. \tag{18}$$

Since for the considered structural and logical model of failure detection and localization we accept that all the diagnostic features for failure detection $\neg\varphi_2(d_{iL,iC}(t))$ at the lowest hierarchy level l_{min} refer to independent failures, then we can form a matrix of markers with independent failures $\lambda(t)$ for hierarchy level l_{min}:

$$\lambda_{iL_{\min},iC}(t) = \neg\varphi_2(d_{iL,iC}(t)). \tag{19}$$

If $l_{min} < l(S_1)$, then it is necessary to define other SCADA hierarchy levels $l_{min} < l \le l(S_1)$. For these hierarchy levels we can detect independent failures when analyzing current diagnostic matrix $D(t)$. It is possible to take into account permissible changes of the controlled parameter state when passing up through SCADA hierarchy levels.

Consider the algorithm of independent failure detection in the event of SCADA low level for failure detection belonging to backbone nodes.

For row iL of diagnostic matrix $D(t)$ the number of diagnostic features $\varphi_0(d_{iL,iC}(t))$ and the number of diagnostic features $\varphi_1(d_{iL,iC}(t))$ can be defined using the following formulas:

$$n_{\varphi_0}(iL,1,i_N,t) = \sum_{iC=1}^{i_N} \varphi_0(d_{iL,iC}(t)), \tag{20}$$

$$n_{\varphi_1}(iL,1,i_N,t) = \sum_{iC=1}^{i_N} \varphi_1(d_{iL,iC}(t)). \tag{21}$$

The total number of diagnostic features for failure detection $\neg\varphi_2(d_{iL,iC}(t))$ for row

iL of diagnostic matrix *D*(*t*) is as follows:

$$n_{\overline{\varphi_2}}(iL,1,i_N,t) = n_{\varphi_0}(iL,1,i_N,t) + n_{\varphi_1}(iL,1,i_N,t). \quad (22)$$

According to the logic of SCADA functioning, at the system's upper hierarchy levels $l_{min} < l_{h+1}$ which refer to backbone nodes, a necessary but insufficient criterion of having independent failures η is an increase in the number of diagnostic features for failure detection compared to the system's lower hierarchy level $l_{min} \le l_h$.

$$\begin{cases} l_h = l_{\min} + 2 \times h \le l(S_1), \quad h = 1,2,\ldots \\ iL_h = l(S_1) + 1 - l_h \\ n\neg\varphi_2(iL_h,1,i_N,t) = n_{\varphi_0}(iL_h,1,i_N,t) + n_{\varphi_1}(iL_h,1,i_N,t) \end{cases}. \quad (23)$$

If at the system's upper hierarchy level l_{h+1}, the number of diagnostic features for failure detection increases compared to the system's lower hierarchy level l_h, we can make a conclusion that the necessary condition of having independent failures was fulfilled at SCADA hierarchy level l_{h+1}.

$$\eta(iL_{h+1}) = (n_0(iL_{h+1},1,i_N,t) - n_0(iL_h,1,i_N,t) > 0) \vee (n_1(iL_{h+1},1,i_N,t) - n_1(iL_h,1,i_N,t) > 0). \quad (24)$$

If $\eta(iL_{h+1}) = 1$, then we can calculate the number of diagnostic features of independent failure detection at SCADA hierarchy level $l_{h+1} \in L_1$.

The function of distinction between independent and secondary failures when controlled parameters pass through SCADA hierarchy levels taking into account the result of transmitting and receiving data process between adjacent hierarchy levels is as follows:

$$f_4(x,y,z) = (2xyz - 2x^2yz - 2xy^2z - xyz^2 + x^2y^2 + 2x^2y^2z + x^2yz^2 + xy^2z^2 - 2x^2y^2z^2) \pmod 3, \quad (25)$$

where x – the controlled parameter state at a transmitting hierarchy level, y – the result of transmitting and receiving data process between adjacent hierarchy levels; z – the controlled parameter state at a receiving hierarchy level; $f_4(x,y,z) = 1$ – a necessary but insufficient condition of having diagnostic features for independent failure detection; $f_4(x,y,z) = 0$ – a sufficient condition of absence of diagnostic features for independent failure detection.

We can define the number of diagnostic features for independent failure detection at hierarchy level $l_{\min} < l_{h+1} \le l(S_1)$ on the basis of formula 24:

$$n_{f_4}(iL,1,i_N,t) = \sum_{iC=1}^{i_N} f_4(d_{iL+2,iC}(t), d_{iL+1,iC}(t), d_{iL,iC}(t)). \quad (26)$$

If $n_{f_4}(iL,1,i_N,t) = 0$, then this criterion is a sufficient condition that no failures have been detected at a current hierarchy level.

If $n_{f_4}(iL,1,i_N,t) > 0$, then this criterion is a sufficient condition of having independent failures.

We can form a row marker matrix of independent failures $\Lambda(t)$ for a current hierarchy level:

$$\lambda_{iL,iC}(t) = f_4(d_{iL+2,iC}(t), d_{iL+1,iC}(t), d_{iL,iC}(t)). \quad (27)$$

Accordingly, predicate $M(i,l)$ of having independent failure markers for controlled parameter x_i at hierarchy level l is as follows:

$$M(i,l) := \lambda_{(S1)+1-l,i}. \quad (28)$$

We can define certain predicates of diagnostic feature distribution through SCADA

structural elements taking into account the characteristic attributes for each hierarchy level. The predicate of existence of at least one controlled parameter x_i at hierarchy level l which has a diagnostic feature for independent failure detection is as follows:

$$\exists i P_1(i,l) := \exists i(M(i,l) \,\&\, \neg S(i,2,l)). \quad (29)$$

The predicate of existence of at least one controlled parameter x_i at hierarchy level l which has a diagnostic feature for independent failure detection with value y is as follows:

$$\exists i P_{11}(i,y,l) := \exists i(M(i,l) \,\&\, S(i,y,l)). \quad (30)$$

The predicate of existence of at least two different controlled parameters x_i and x_j at hierarchy level l which have a diagnostic feature for independent failure detection is as follows:

$$\exists i \exists j P_2(i,j,l) := \exists i \exists j((i \neq j) \,\&\, P_1(i,l) \,\&\, P_1(j,l)). \quad (31)$$

The predicate of existence of at least two controlled parameters x_i and x_j at hierarchy level l which have different diagnostic features for failure detection is as follows:

$$\exists i \exists j P_3(i,j,l) := \exists i \exists j((i \neq j) \,\&\, M(i,l) \,\&\, M(j,l) \,\&\, ((S(i,0,1) \,\&\, S(j,1,l)) \vee (S(i,1,l) \,\&\, S(j,0,l)))). \quad (32)$$

Thus, for the structural and logical model we consider, the number of independent failures $n_F(l=2 \vee l=3)$ for controlled parameters with timestamp t at hierarchy level Sensors/Ch1 can be defined as follows:

– We verify whether there are at least two different controlled parameters x_i and x_j which have diagnostic features of independent failure detection at hierarchy level Sensors/Ch1. Then we verify there are no Sensors/Ch1 for which we have at least two controlled parameters x_i and x_j having different diagnostic features for independent failure detection. This means that the number of independent failures at a current hierarchy level is equal to the number of hierarchy modules for which we have at least one diagnostic feature of independent failure detection;
– We verify whether there is at least one Sensors/Ch1 at hierarchy level Sensors/Ch1 for which there are at least two controlled parameters x_i and x_j having different diagnostic features for independent failure detection. This means that the number of independent failures at a current hierarchy level is calculated by the number of different diagnostic features for independent failure detection per structural module.

The number of independent failures for the levels of SCADA hierarchy is determined in analogy to hierarchy level Sensors/Ch1. The foregoing is achieved by taking into account the connection between various hierarchy levels.

Conclusions

The system of predicates we considered can be applied when forming a knowledge base of an expert diagnostic system. It allows us to implement a method for SCADA failure diagnostics. It takes into account the consistencies of information flow changes in real time in the event of incomplete / unreliable / absent data in the system's structural elements. This method of independent failure detection and localization ensures the reliability of SCADA operational monitoring.

References

1. Giarratano J. Expert Systems: Principles and Programming / J. Giarratano, G. Riley. – [4th Edition]. – Course Technology, 2004. – P. 842.
2. Varlamov O. Practical Guide on Creation of Miwitary Expert Systems / O. Varlamov, M. Chibirova, A. Khadiev, P. Antonov, G. Sergushin, I. Shoshev, K. Nazarov. – Tutorial. – M: NII MIVAR, 2016. – P. 184.
3. Ruchkin V. Universal Artificial Intelligence and Expert Systems / V. Ruchkin, V. Fulin. – St. Petersburg: BHV-Peterburg, 2009. – P. 240.

4. Syrotkina O. Software Diagnostics for Reliability of SCADA Structural Elements / O. Syrotkina, M. Alekseyev // Power Engineering and Information Technologies in Technical Objects Controls: Taylor & Francis Group, London. – 2016. – P. 259–265.
5. Syrotkina O. Automatic Diagnosis Method for SCADA Operability / O. Syrotkina // Quality Control Tools and Techniques. – Ivano-Frankivsk, 2015. – V. 1. – P. 19–26.

РЕЗЮМЕ

М. Алексєєв, І. Удовик, О. Сироткіна

Застосування логіки предикатів для виявлення відмов у SCADA системах

У даній статті розглядається задача виявлення та локалізації відмов у SCADA в режимі реального часу на основі аналізу зміни стану інформаційних потоків системи у процесі їх проходження за структурними елементами та рівнями ієрархії. Великий обсяг та інтенсивний потік низькорівневої діагностичної інформації, що генерується SCADA, вимагає розробки універсальної та ефективної системи подання знань стосовно до експертної діагностичної системи підтримки прийняття рішень. Розглядається розроблена предикатна система подання знань, перевагами якої є простота реалізації та універсальність опису задачі.

Пропонується структурно-логічна модель для опису SCADA системи будь-якої топології. В рамках даної моделі визначаються предикати наявності зв'язку між структурними елементами системи різних рівнів ієрархії.

На основі характеристичних функцій тризначної логіки формуються необхідні та достатні діагностичні ознаки виявлення / відсутності відмови у системі, розмежування незалежних і вторинних відмов.

Визначається предикатна система подання знань для реалізації методу діагностики працездатності SCADA в умовах неповних даних / недостовірних даних. Ефективний алгоритм пошуку рішення на основі запропонованої системи предикатів дозволяє проводити оперативний контроль стану структурних елементів SCADA.

Надійшла до редакції 31.10.2017

