

Wyznaczanie struktur diagnozowania porównawczego przy użyciu algorytmu genetycznego

Łukasz STRZELECKI

Instytut Teleinformatyki i Automatyki, Wydział Cybernetyki, WAT,
ul. Gen. S. Kaliskiego 2, 00-908 Warszawa
lukasz.strzelecki@wat.edu.pl

STRESZCZENIE: W artykule wykazano, że metoda adaptacyjna – algorytm genetyczny – może być skutecznie wykorzystana do wyznaczania struktur diagnozowania porównawczego. Zaproponowano odpowiednią funkcję przystosowania, a także przeprowadzono analizę jej użyteczności w oparciu o wyniki badań empirycznych.

SŁOWA KLUCZOWE: diagnostyka systemowa, struktury diagnozowania porównawczego, model MM, algorytm genetyczny, metoda adaptacyjna

1. Wprowadzenie

W systemach rozproszonych uzyskanie wysokiej wiarygodności wyniku przetwarzania danych jest możliwe m.in. poprzez zastosowanie metod diagnostyki systemowej do wykrywania uszkodzonych węzłów, których wykorzystanie mogłoby prowadzić do uzyskania błędnych rezultatów obliczeń. Jedną z metod diagnostyki systemowej jest metoda prób porównawczych MM [7], w której wnioskowanie o stanie niezawodnościowym sieci odbywa się na podstawie wzajemnego testowania się wybranych węzłów.

W praktyce wykonanie testu obarczone jest pewnym *uogólnionym kosztem*¹. Zatem ekonomicznie uzasadnione jest poszukiwanie takich *struktur diagnozowania porównawczego* (SDP), które implikowałyby będą wykonanie minimalnej liczby testów koniecznych do zidentyfikowania niezdalnych węzłów

¹ Np. w sensie czasu i/lub energii potrzebnych do jego realizacji.

i z tego powodu będą najtańsze. Jakkolwiek znane są analityczne metody wyznaczania struktur najtańszych, to są one skomplikowane i/lub dosyć kosztowne obliczeniowo [9]. Natomiast wielokrotnie wystarczającym jest wyznaczenie struktury *quasi*-najtańszej, jednakże spełniającej określone wymagania (m.in. umożliwiającej zidentyfikowanie wszystkich niezdatnych węzłów, pod warunkiem, że ich łączna liczba jest nie większa, niż t). Sytuacja ta sugeruje, że metoda adaptacyjna, którą jest *algorytm genetyczny* [2], może stanowić użyteczne narzędzie wyznaczania tego typu struktur diagnozowania porównawczego, co jest przedmiotem tego artykułu.

Artykuł został podzielony na 6 części, z których pierwszą jest niniejsze wprowadzenie. W drugiej części została podana definicja *struktury diagnozowania porównawczego* oraz zostały przytoczone niektóre jej własności. W trzeciej części zaproponowano (nowy) sposób reprezentacji zbioru prób porównawczych struktury diagnozowania porównawczego, który umożliwia jej przetwarzanie z wykorzystaniem algorytmu genetycznego. W czwartej części zaproponowano funkcję przystosowania dla algorytmu genetycznego umożliwiającego wyznaczenie *quasi*-najtańszych SDP o określonych cechach charakterystycznych, zaś w piątej części poddano weryfikacji jej skuteczność. W podsumowaniu (część szósta) odniesiono się do uzyskanych rezultatów, a także wskazano kierunek dalszych badań.

2. Podstawowe określenia i własności

Niech spójny graf zwykły G ($E(G)$ - zbiór węzłów, $U(G)$ - zbiór krawędzi) reprezentuje strukturę logiczną sieci (np. komputerów). Struktura ta jest *t*-diagnozowalna metodą prób porównawczych (MM) jeżeli umożliwia zlokalizowanie wszystkich niezdatnych węzłów pod warunkiem, że jest ich nie więcej, niż t .

Niech $\Psi(G)$ oznacza zbiór prób porównawczych dla sieci o strukturze logicznej G , $\psi \in \Psi(G)$, $K(\psi)$ - komparator próby ψ , a $P(\psi)$ - parę porównawczą próby ψ . Mówimy, że zbiór $\Psi' \subseteq \Psi(G)$ jest *pokryciem zbioru węzłów struktury logicznej G* , jeżeli $\bigcup_{\psi \in \Psi'} P(\psi) = E(G)$.

Strukturą diagnozowania porównawczego sieci G nazywamy taką dwójkę uporządkowaną $\langle G, \Psi' \rangle$ ($\Psi' \subseteq \Psi(G)$), że zbiór prób porównawczych Ψ' jest pokryciem struktury logicznej G , czyli $\{P(\psi) : \psi \in \Psi'\} = E(G)$.

Grafem porównań struktury diagnozowania porównawczego $\langle G, \Psi' \rangle$ ($\Psi' \subseteq \Psi(G)$) nazywamy taki graf zwykły $\hat{G}(G, \Psi') = \langle E(G), U(G, \Psi') \rangle$

o opisanych krawędziach, że $[(e', e'') \in U(G, \Psi')] \Leftrightarrow [\exists \psi \in \Psi': P(\psi) = \{e', e''\}]$, w którym etykieta krawędzi (e', e'') jest $K(\psi)$.

Sieć o strukturze logicznej G nazywana jest *jednokroowo t -diagnozowalną* za pomocą zbioru prób porównawczych $\Psi' \subseteq \Psi(G)$, jeżeli każda para takich zbiorów $E' \subseteq E(G)$ i $E'' \subseteq E(G)$ niezdatnych węzłów, że $|E'| \leq t$ i $|E''| \leq t$, jest rozróżnialna za pomocą (choć jednej) próby porównawczej $\psi \in \Psi'$.

Warunkiem koniecznym [6], aby struktura logiczna G była *t -diagnozowalna* metodą prób porównawczych jest, aby

$$\left(|E(G)| \geq 2t + 1\right) \wedge \left(|E(e)| \geq t \ (e \in E)\right) \quad (1)$$

gdzie $E(e)$ - zbiór węzłów przyległych do węzła e .

Zgodnie z [8] warunkiem wystarczającym, by struktura logiczna G była *t -diagnozowalna* za pomocą zbioru prób porównawczych $\Psi' \subseteq \Psi(G)$ jest, aby

$$\forall s \in \{0, \dots, t-1\} \forall E' \subseteq E(G) : |E'| = |E(G)| - 2t + s : |T(E')| > s, \quad (2)$$

gdzie

$$T(E') = \left\{ e \in E(G) \setminus E' : \exists \psi \in \Psi' : |P(\psi) \cap \{e\}| = |K(\psi) \cap E'| = |P(\psi) \cap E'| = 1 \right\}.$$

3. Transformacja struktury diagnozowania porównawczego do postaci wektora binarnego

Klasyczny algorytm genetyczny operuje na wektorach binarnych reprezentujących konkretne rozwiązania zadanego problemu. Zatem jego wykorzystanie do wyznaczania najtańszej struktury *t -diagnozowalnej* metodą prób porównawczych wymaga opracowania sposobu przedstawiania struktur diagnozowania porównawczego w formie wektorów binarnych.

Węzłom grafu G można przypisać etykiety od 1 do $|E(G)|$. W takim przypadku zbiór prób porównawczych $\Psi' \subseteq \Psi(G)$ może zostać przedstawiony, jako trójwymiarowa macierz binarna $M_{|E(G)| \times |E(G)| \times |E(G)|}$, w której:

$$m_{ijk} = 1 \rightarrow m_{jik} = 1 \rightarrow \exists \psi \in \Psi' : (K(\psi) = \{e_k\} \wedge P(\psi) = \{e_i, e_j\}) \quad (3)$$

gdzie:

m_{ijk}, m_{jik} – elementy macierzy M ,

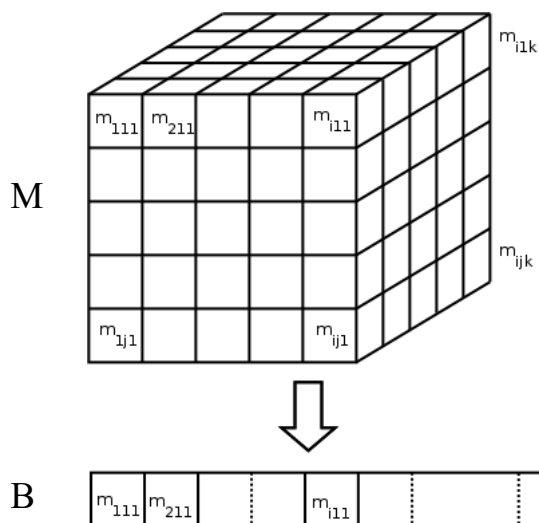
e_i, e_j, e_k – węzły grafu G o etykietach i, j oraz k .

Macierz M może następnie zostać sprowadzona do postaci wektora binarnego B (o długości $|E(G)|^3$) poprzez połączenie jej elementów w taki sposób, że

$$m_{ijk} = b_{(k-1) \cdot i \cdot j + (j-1) \cdot i + i} \quad (4)$$

gdzie: b_x jest elementem wektora binarnego B zlokalizowanym na pozycji x .

Przykład opisanego przekształcenia został zobrazowany na rysunku 1.



Rys. 1. Transformacja przykładowej macierzy M do postaci wektora binarnego B

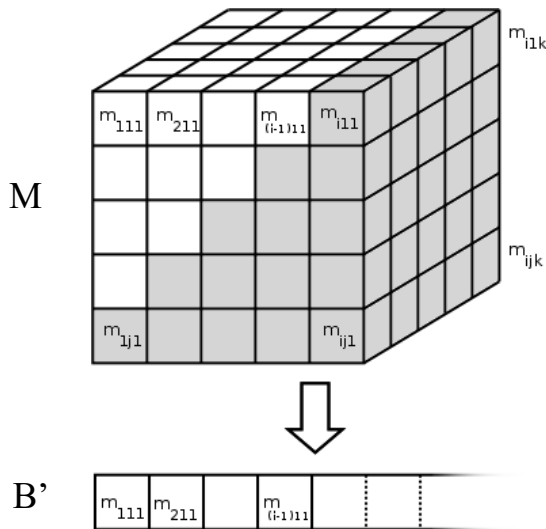
Złożoność pamięciowa algorytmu genetycznego jest uzależniona od liczby i długości chromosomów, na których operuje. W tym kontekście warto uwzględnić, że macierz M jest symetryczna, gdyż $(m_{ijk} = 1 \rightarrow m_{jik} = 1) \wedge (m_{ijk} = 0 \rightarrow m_{jik} = 0)$. Ponadto $i \neq j \wedge i \neq k \wedge j \neq k$, co wynika bezpośrednio z definicji próby porównawczej. Zatem macierz M można sprowadzić do takiego wektora binarnego B' , że:

$$m_{ijk} = b'_{(|E(G)|^2 - |E(G)|) \cdot (k-1) + \frac{2 \cdot (|E(G)| - 1) - (j-2)}{2} \cdot (j-1) + i} \quad (5)$$

gdzie b_x jest elementem wektora binarnego B' zlokalizowanym na pozycji x .

Działanie to pozwala ograniczyć długość chromosomów (reprezentujących zbiory prób porównawczych) z $|E(G)|^3$ do $\frac{|E(G)|^3 - |E(G)|^2}{2}$, czyli ponad

dwukrotnie. Przykład opisanego przekształcenia został pokazany na rysunku 2.



Rys. 2. Transformacja przykładowej macierzy M do postaci wektora binarnego B' (elementy oznaczone kolorem szarym nie są uwzględniane w wektorze B')

4. Konstrukcja funkcji przystosowania

Z warunków koniecznego (1) i wystarczającego wyniku (2), że dla t -diagnozowalnej struktury diagnozowania porównawczego zachodzi zależność $\forall e \in E(G) |\{\psi \in \Psi' : e \in P(\psi)\}| \geq t$, zaś dla macierzy M odpowiadającej zbiorowi prób porównawczych zachodzi zależność $\forall_{\{i,j\} \in \{1, \dots, |E(G)|\}} \sum_{k=1}^{|E(G)|} m_{ijk} \geq t$. Zwróćmy uwagę, że uogólniony koszt struktury diagnozowania (tj. łączny uogólniony koszt testów przewidywanych przez zbiór prób porównawczych) maleje wraz ze zbliżaniem się uzyskiwanych wartości (w podanej zależności) do t , gdyż oznacza to zmniejszanie liczby wykonywanych testów.

Rozsądnym jest (m.in. ze względów praktycznych), aby realizacja prób porównawczych była równomiernie rozłożona między węzły mogące pełnić rolę komparatorów², co można sprowadzić do wymagania, aby wartość $\forall_{\{i,k\} \in \{1, \dots, |E(G)|\}} \sum_{j=1}^{|E(G)|} m_{ijk}$ była możliwie najmniejsza.

Uwzględniając powyższe oraz warunek (2) można zaproponować następującą funkcję przystosowania:

² Węzły posiadające tylko jeden węzeł przyległy nie mogą być komparatorami.

$$f((G, \Psi')) = \sum_{e \in E(G)} g_1(e, \Psi') + \sum_{e \in \{K(\psi) : \psi \in \Psi'\}} g_2(e, \Psi') + \quad (6)$$

$$+ \sum_{s=0}^{t-1} \sum_{E' \subseteq E(G) : |E'| = |E(G) - 2 \cdot t + s|} g_3(s, E')$$

gdzie:

- $g_1(e, \Psi') = \begin{cases} t & \text{dla } k(e, \Psi') = t \\ -|t - k(e, \Psi')| & \text{w p.p.} \end{cases}$
gdzie $k(e, \Psi') = |\{\psi \in \Psi' : e \in P(\psi)\}|$

- $g_2(e, \Psi') = \begin{cases} t & \text{dla } z(e, \Psi') = t \\ -|t - z(e, \Psi')| & \text{w p.p.} \end{cases}$
gdzie $z(e, \Psi') = |\{\psi \in \Psi' : K(e) = \{e\}\}|$

- $g_3(s, E') = \begin{cases} t & \text{dla } |T(E')| > s \\ -|s - |T(E')|| & \text{w p.p.} \end{cases}$
gdzie $T(E')$ jest zbiorem zdefiniowanym w (2).

Pierwszy człon funkcji przystosowania ($\sum_{e \in E(G)} g_1(e, \Psi')$) nakierowuje algorytm genetyczny na poszukiwanie takich rozwiązań³, w których każdy węzeł jest elementem t par porównawczych, co ma na celu spełnienie warunku koniecznego przez odnajdowane rozwiązania.

Drugi człon funkcji przystosowania ($\sum_{e \in \{K(\psi) : \psi \in \Psi'\}} g_2(e, \Psi')$) promuje takie struktury diagnozowania porównawczego, w których każdy komparator jest elementem, co najwyżej, t prób porównawczych. Restrykcja ta wymusza poszukiwanie rozwiązań o możliwie najmniejszej liczbie prób porównawczych, a zatem także najtańszych. Dodatkową zaletą omawianego członu jest wymuszenie równomiernego rozłożenia prób porównawczych między wszystkie komparatory.

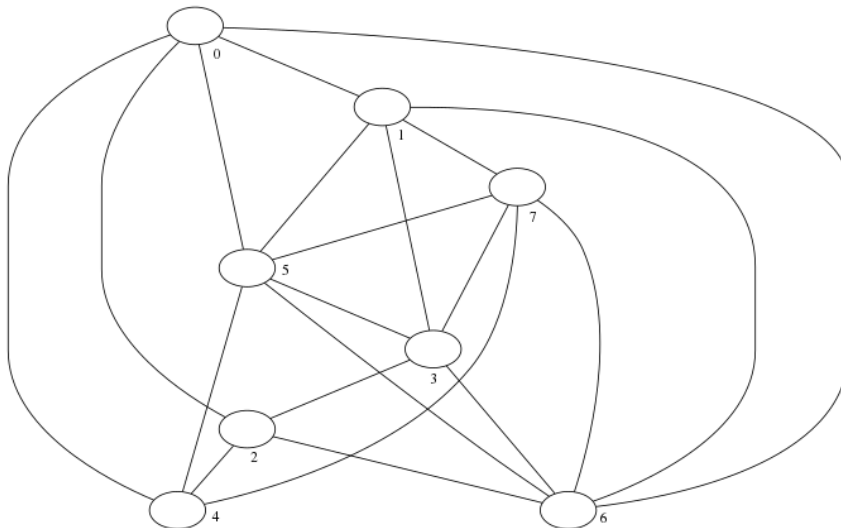
Wykorzystanie trzeciego członu funkcji przystosowania ($\sum_{s=0}^{t-1} \sum_{E' \subseteq E(G) : |E'| = |E(G) - 2 \cdot t + s|} g_3(s, E')$) wynika bezpośrednio z wymogu poszukiwania takich struktur diagnozowania porównawczego, dla których spełniony jest warunek (2), czyli takich, które są t -diagnozowalne.

³ Chromosomów odpowiadających trójwymiarowym macierzom binarnym reprezentującym zbiory prób porównawczych dla struktury logicznej G .

5. Ocena efektywności zaproponowanej metody

W celu praktycznej weryfikacji przydatności zaproponowanej adaptacyjnej metody wyznaczania najtańszych struktur diagnozowania porównawczego, zaimplementowano algorytm genetyczny oraz opracowaną funkcję przystosowania. Przeprowadzone testy empiryczne (obejmujące ponad 10 000 przypadków dla różnych wartości t) pozwoliły na ustalenie, że skuteczność⁴ rozpatrywanej metody sięga 96%, a więc wyznaczone SDP są w zdecydowanej większości (*quasi-*) najtańsze i *t*-diagnozowalne. Warto dodać, że analiza 4% wyznaczonych SDP, które nie spełniały wymagań (nie były *t*-diagnozowalne) wykazała, że są one *quasi-t*-diagnozowalne, tj. ich niewielka modyfikacja (polegająca na dodaniu do zbioru Ψ' nie więcej, niż $2t$ odpowiednich prób porównawczych) umożliwia przekształcenie ich w SDP *t*-diagnozowalne⁵.

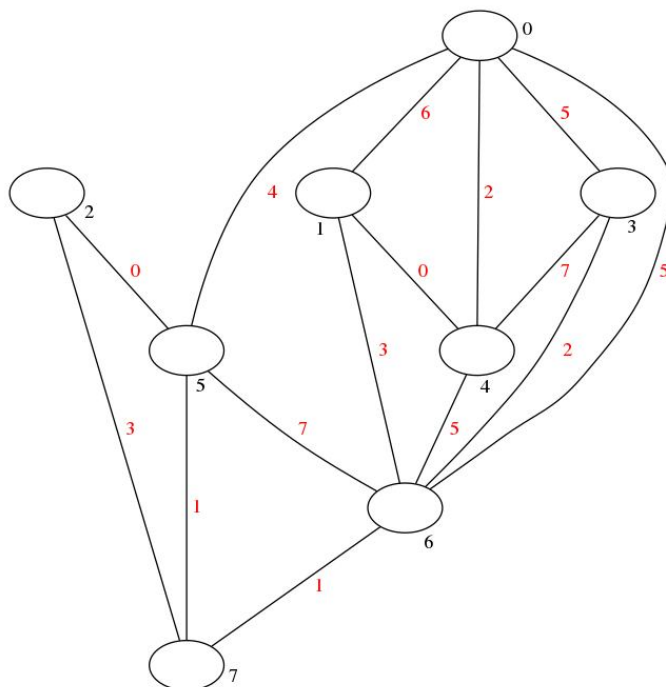
W celach poglądowych na rysunku numer 3 został przedstawiony obraz struktury logicznej G , dla której (przy wykorzystaniu opracowanej metody) została wyznaczona *t*-diagnozowalna SDP, której graf porównań został zobrazowany na rysunku 4.



Rys. 3. Obraz przykładowej struktury logicznej G

⁴ Skuteczność jest rozumiana, jako częstość zdarzenia, że wyznaczona struktura diagnozowania porównawczego będzie *quasi*-najtańsza i *t*-diagnozowalna.

⁵ Analiza została wykonana przy wykorzystaniu narzędzia programowanego (specjalnie zaimplementowanego w tym celu) umożliwiającego określenie minimalnej liczby prób porównawczych, które należy dodać do zbioru $\Psi' \subseteq \Psi(G)$, aby struktura $\langle G, \Psi' \rangle$ była *t*-diagnozowalna.

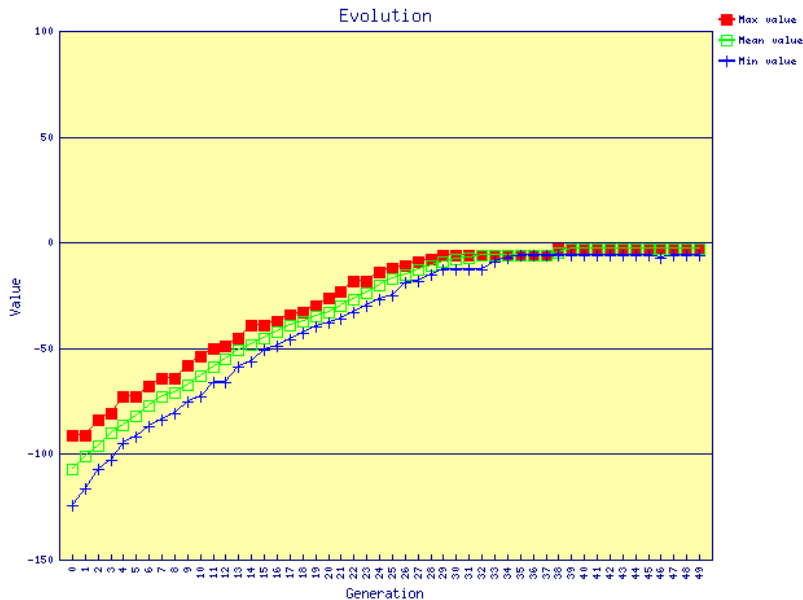


**Rys. 4. Obraz grafu porównań SDP $\langle G, \Psi' \rangle (\Psi' \subseteq \Psi(G))$
(kolorem czerwonym zostały oznaczone etykiety komparatorów)**

Na rysunku 5 zamieszczony został wykres przedstawiający ewolucję kolejnych generacji populacji (chromosomów przetwarzanych przez algorytm genetyczny) obserwowaną podczas wyznaczania zbioru prób porównawczych dla struktury logicznej G zobrazonej na rysunku 3. Łatwo zauważyć dosyć szybki (i zbieżny) wzrost wartości oceny rozwiązań najlepszych, najgorszych, jak i średniej oceny całej populacji, co jednoznacznie potwierdza skuteczność zaproponowanej metody. Wykorzystane parametry algorytmu genetycznego zostały zestawione w tabeli 1.

Tab. 1. Parametry algorytmu genetycznego

| Parametr | Wartość |
|--------------------------------|---------------|
| Populacja | 100 |
| Mutacja | 0.05 |
| Metoda selekcji | ruletka |
| Krzyżowanie | jednopunktowe |
| Prawdopodobieństwo krzyżowania | 0.9 |



Rys. 5 Zobrazowanie rozwoju populacji

W trakcie dalszej części badań poddano ocenie przydatność poszczególnych członów opracowanej funkcji przystosowania. Na podstawie wyników przeprowadzonych testów empirycznych (obejmujących 10 000 przypadków) ustalono, że brak pierwszego i/lub drugiego członu obniża skuteczność funkcji przystosowania do nieakceptowalnego poziomu poniżej 40%. Natomiast pominięcie trzeciego członu zmniejsza jej skuteczność do poziomu 83%, jednakże istotnie zwiększa efektywność (rozumianą, jako złożoność obliczeniową oraz pamięciową)⁶. Na tej podstawie można przyjąć, że w niektórych zastosowaniach (uwarunkowanych czasowo) zasadne może być wykorzystanie funkcji przystosowania w postaci:

$$f(\langle G, \Psi' \rangle) = \sum_{e \in E(G)} g_1(e, \Psi') + \sum_{e \in \{K(\psi) : \psi \in \Psi'\}} g_2(e, \Psi') \quad (7)$$

W takim przypadku można rozważyć kilkukrotne uruchomienie algorytmu genetycznego dla uproszczonej funkcji przystosowania (do momentu uzyskania satysfakcjonującego wyniku) zamiast jednokrotnego uruchamiania algorytmu genetycznego dla jej postaci oryginalnej (składającej się z trzech członów).

⁶ Podany sposób rozumienia efektywności funkcji przystosowania przekłada się wprost na jej złożoność czasową.

6. Podsumowanie

Celem artykułu było wykazanie, iż metoda adaptacyjna (algorytm genetyczny) może zostać zastosowana do wyznaczania najtańszych (w sensie uogólnionego kosztu) *t-diagnozowalnych* struktur diagnozowania porównawczego. W artykule (w części 3) zaproponowano macierzowy sposób zapisu informacji o zbiorze prób porównawczych umożliwiającą jego (dalszą) konwersję do postaci wektora binarnego, który może stanowić chromosom w algorytmie genetycznym. W części 4 zaproponowano funkcję przystosowania skonstruowaną w oparciu o warunki konieczny i wystarczający dla struktur *t-diagnozowalnych* oraz uwzględniającą ich specyficzne cechy. Część piąta została poświęcona omówieniu wyników empirycznych testów mających na celu określenie skuteczności zaproponowanej metody wyznaczania najtańszych struktur diagnozowania porównawczego.

Uzyskane wyniki jednoznacznie potwierdzają przydatność opracowanego rozwiązania – struktury najtańsze są odnajdowane przez algorytm genetyczny w czasie dopuszczalnym dla większości systemów teleinformatycznych. Jednakże należy pamiętać, że w pracy przyjęto założenie, iż uogólniony koszt wszystkich prób porównawczych jest jednakowy oraz istnieją łącza komunikacyjne pomiędzy wszystkimi węzłami struktury logicznej. Jakkolwiek w przypadku systemów pracujących w sieci Internet nie jest to ograniczeniem, to w innych sytuacjach może uniemożliwić praktyczne wykorzystanie zaproponowanej metody. Rozwiązanie tego problemu będzie stanowiło kolejny etap badań.

Literatura

- [1] ARCIUCH A.: *Techniczne aspekty diagnozowania sieci procesorów o łagodnej degradacji typu sześcián 4-wymiarowy metodą prób porównawczych*. Przegląd Teleinformatyczny, Nr 2 / 2013, s. 3-11.
- [2] GOLDBERG D. E.: *Algorytmy genetyczne i ich zastosowania*. WNT, Warszawa, 2003.
- [3] KULESZA R., ZIELIŃSKI Z.: *Wnikliwość diagnozowania sieci procesorów metodą porównawczą*. W. *Systemy czasu rzeczywistego. Postępy badań i zastosowania*. WKŁ, Warszawa, 2009, s. 199-210.
- [4] KULESZA R., ZIELIŃSKI Z.: *Diagnosis resolution of processors' network using the comparison method*. Przegląd Elektrotechniczny (Electrical Review), No. 9/2010, s. 157-162.

- [5] KULESZA R., ZIELIŃSKI Z.: *The life period of the hypercube processors' network diagnosed with the use of the comparison method*. Monographs On System Dependability, Oficyna Wydawnicza Politechniki Wrocławskiej, 2010.
- [6] MALEK M.: *A Comparison Connection Assignment for Diagnosis of Multiprocessor Systems*. Proc. Seventh Int'l Symp. Computer Architecture, 1980, pp. 31-35.
- [7] MAENG J., MALEK M.: *A Comparison Connection Assignment for Self-Diagnosis of Multiprocessor Systems*. Digest Int'l Symp. FTC, 1981, pp. 173-175.
- [8] SENGUPTA A., DAHBURA A.T.: *On Self-Diagnosable Multiprocessor Systems: Diagnosis by the Comparison Approach*. IEEE Trans. Comput., vol. 41, no. 11, Nov. 1992, pp. 1386-1396.
- [9] STRZELECKI Ł.: *Metody projektowania najtańszych t-diagnozowalnych struktur diagnostyki systemowej dla sieci procesorów typu binarnego sześcianu 4-wymiarowego*. Rozprawa doktorska. Wojskowa Akademia Techniczna, Warszawa, 2012.
- [10] ZIELIŃSKI Z.: *Podstawy diagnostyki systemowej sieci procesorów o łagodnej degradacji i strukturze hipersześcianu*. Wojskowa Akademia Techniczna, Warszawa, 2012.

Design of comparison structures using genetic algorithm

ABSTRACT: In the article was proven, that adaptive method - genetic algorithm - can be used during designing of comparison structures. There was proposed suitable adaptation function and there were presented results of empirical experiments, that confirm usability of developed solution.

KEYWORDS: system level diagnosis, MM model, genetic algorithm

Praca wpłynęła do redakcji: 28.12.2016 r.