

Projektowanie struktur opiniowania diagnostycznego z wykorzystaniem algorytmu genetycznego

Łukasz STRZELECKI, Zbigniew ZIELIŃSKI

Zakład Teleinformatyki, Instytut Teleinformatyki i Automatyki WAT,
ul. Kaliskiego 2, 00-908 Warszawa

STRESZCZENIE: W artykule rozpatrzono problemy występujące przy projektowaniu struktur opiniowania diagnostycznego (OD) dla sieci komputerowych o ustalonej strukturze logicznej (topologii). Przedstawiono adaptacyjną metodę wykorzystującą algorytm genetyczny, która umożliwia projektowanie najtańszych struktur OD dla danej topologii.

SŁOWA KLUCZOWE: diagnostyka systemowa, struktury opiniowania diagnostycznego, modele PMC i BGM, algorytm genetyczny, projektowanie

1. Wprowadzenie

Koniecznym warunkiem uzyskania wysokiej efektywności procesu identyfikacji niezdatnych węzłów (komputerów) sieci teleinformatycznej w przypadku stosowania metod diagnostyki systemowej [3, 5] jest odpowiednio zaprojektowana struktura diagnostyczna. Jedną z bardziej znanych metod diagnostyki systemowej jest metoda opiniowania diagnostycznego [3, 4, 7, 8]. Wnioskowanie o stanie niezawodnościowym sieci odbywa się w tym przypadku na podstawie wyników wszystkich testowań, określonych przez strukturę diagnostyczną.

Strukturą opiniowania diagnostycznego (OD) systemu nazywamy spójny digraf (unigraf zorientowany) $G = \langle E, U \rangle$ bez pętli ($\langle e', e' \rangle \notin U, e' \in E$), w którym łuk $\langle e', e'' \rangle$ oznacza, że węzeł e' sieci wyraża opinię (na podstawie wyniku testowania) o stanie niezawodnościowym węzła e'' . Strukturę umożliwiającą

identyfikację m niezdatnych elementów i zawierającą minimalną liczbę łuków nazywamy optymalną strukturą m -diagnozowalną.

W wielu sytuacjach do racjonalnego doboru struktury konieczne jest uwzględnienie pewnych uogólnionych kosztów wzajemnego testowania się elementów sieci teleinformatycznej, których w ogólnym przypadku nie można uważać za jednakowe. Wówczas możemy mówić o potrzebie wyznaczenia struktury najtańszej, tj. takiej, dla której sumaryczny uogólniony koszt testowania przyjmuje wartość minimalną.

W literaturze znane są metody analityczne wyznaczania struktur OD typu PMC i BGM [1, 2, 4]. W pracach [7, 8] została przedstawiona adaptacyjna metoda wyznaczania najtańszych struktur OD. Zaproponowany w [8] adaptacyjny algorytm bazuje na dekompozycji zadania na dwa sekwencyjnie wykonywane podzadania: wyznaczenie składowej silnej spójności oraz wyznaczenie części acyklicznej struktury. Ograniczeniem przedstawionego w [8] algorytmu jest operowanie na kosztach wzajemnego testowania się jednostek, określonych dla grafu pełnego.

W praktyce eksploatacji sieci komputerowych niezbędne jest wyznaczenie najtańszej struktury OD dla ustalonej jej struktury logicznej sieci (topologii), co może mieć miejsce po każdej rekonfiguracji sieci, dokonywanej po wykryciu niezdatności określonych węzłów sieci. W artykule przedstawiono adaptacyjną metodę projektowania struktur OD z uwzględnieniem topologii sieci. Dodatkowo zaproponowano uogólnienie metody, umożliwiające wyznaczenie zarówno struktur typów PMC, jak i BGM.

2. Podstawowe określenia

W celu uściślenia terminologii poniżej zostały przedstawione podstawowe określenia wykorzystywane w dalszej części artykułu.

Określenie 1.

Dla m -diagnozowalnej struktury OD typu PMC spełnione są warunki:

1. Warunek konieczny [5]:

$$\left(|E| \geq 2 \cdot m + 1\right) \wedge \left(\forall e \in E : \mu^-(e) \geq m\right), \quad (1)$$

gdzie: $|E|$ oznacza liczebność zbioru E , a $\mu^-(e)$ – stopień wejściowy węzła e , czyli liczbę elementów, które testują węzeł e .

2. Warunek wystarczający [5, 10]:

$$\left((\forall 0 \leq p \leq m-1) \wedge (\forall E' \subset E : |E'| = |E| - 2 \cdot m + p) \right) : |\Gamma(E')| > p, \quad (2)$$

gdzie: $\Gamma(E')$ oznacza zbiór następników E' takich, że $\Gamma(E') \cap E' = \emptyset$.

Określenie 2.

Dla m -diagnozowalnej struktury OD typu BGM spełnione są poniższe warunki:

1. Warunek konieczny [5]:

$$(|E| \geq m + 2) \wedge (\forall e \in E : \mu^-(e) \geq m), \quad (3)$$

gdzie: $\mu^-(e)$ oznacza stopień wejściowy węzła $e \in E$.

2. Warunek wystarczający [3, 5]:

$$\left((\forall e', e'' \in E : \mu^-(e') = \mu^-(e'') = m) \wedge (e' \in \Gamma^{-1}(e'') \cap \Gamma(e'')) \right) \\ \left(\exists e^* \in \Gamma^{-1}(e') \setminus \Gamma^{-1}(e') \cap \Gamma^{-1}(e'') : \Gamma^{-1}(e^*) \neq \Gamma^{-1}(e'') \right), \quad (4)$$

gdzie: $\Gamma(e)$ oznacza zbiór następników węzła e , a $\Gamma^{-1}(e)$ oznacza zbiór poprzedników węzła e .

Klasa struktur BGM obejmuje również struktury klasy PMC¹. Jednakże w rozpatrywanym przypadku jako struktury BGM rozumiane będą tylko takie struktury OD, które spełniają warunki konieczny i wystarczający dla klasy struktur BGM, lecz nie są strukturami typu PMC. W związku z powyższym prawidłowy jest poniższy zapis wcześniej podanego warunku wystarczającego:

$$(\forall e', e'' \in E : \mu^-(e') = \mu^-(e'') = m) : (\Gamma^{-1}(e') \setminus \{e''\} \neq \Gamma^{-1}(e'') \setminus \{e'\}). \quad (5)$$

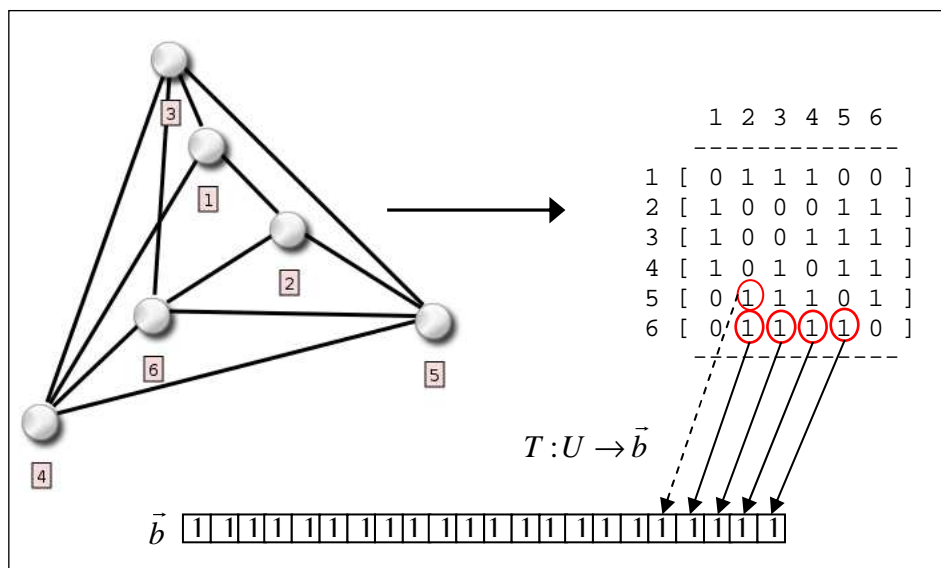
3. Reprezentacja struktury OD

Ogólna zasada działania algorytmu genetycznego (AG) zastosowanego do wspomaganego generowania struktur opiniowania diagnostycznego zarówno typu PMC, jak i BGM została omówiona w [7]. Jednakże zaprezentowany tam algorytm operował na strukturach OD, zapisanych w postaci wektorów bitowych, bez uwzględniania dodatkowych ograniczeń w postaci istniejących połączeń komunikacyjnych. Jednym ze sposobów zdeterminowania metody heurystycznej, jaką jest AG, do poszukiwania rozwiązań określonego typu (w tym przypadku struktur bazujących na grafie o określonych krawędziach),

¹ Każda struktura opiniowania diagnostycznego typu PMC jest również poprawną strukturą opiniowania diagnostycznego typu BGM.

jest takie zdefiniowanie chromosomów (wektorów bitowych), by odnosiły się jedynie do dopuszczalnych rozwiązań (struktur reprezentujących grafy zgodne z zadaną strukturą komunikacyjną).

Założymy, że logiczna struktura sieci (węzłów i możliwych linii komunikacyjnych między węzłami) będzie opisana w postaci grafu niezorientowanego $G_N = \langle E, K \rangle$, gdzie K jest zbiorem krawędzi reprezentujących możliwe ścieżki przesyłania danych między węzłami ($K \subseteq \{\{e', e''\} : e', e'' \in E\}$). Graf $G_N = \langle E, K \rangle$ będzie przedstawiany w postaci macierzy sąsiedztwa M_S . Otrzymaną macierz należy przekształcić do równoważnej macierzy przejść M_P , a następnie zdefiniować bijekcję $T: U \rightarrow \vec{b}$, przekształcającą zbiór łuków, opisanych w macierzy M_P , w elementy wektora bitowego \vec{b} . Przykład takiego przekształcenia został przedstawiony na rysunku 1.



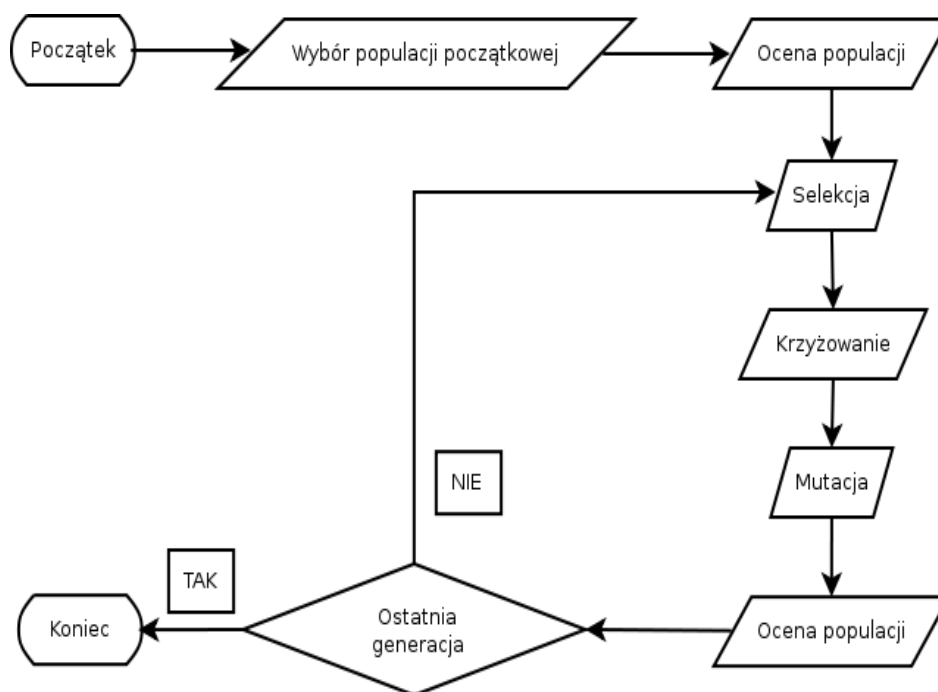
Rys. 1. Przekształcenie zbioru łuków do wektora bitowego

Uzyskany w ten sposób fenotyp należy potraktować jako punkt odniesienia dla chromosomów przetwarzanych przez algorytm genetyczny, który zostanie wykorzystany do projektowania struktury opiniowania diagnostycznego. Oznacza to, że wszystkie chromosomy w przetwarzanej przez AG populacji

będą posiadały taką samą liczbę genów jak wektor bitowy \vec{b} , a jedyne zmiany mogą dotyczyć poszczególnych jego pozycji.

4. Ocena rozwiązań cząstkowych

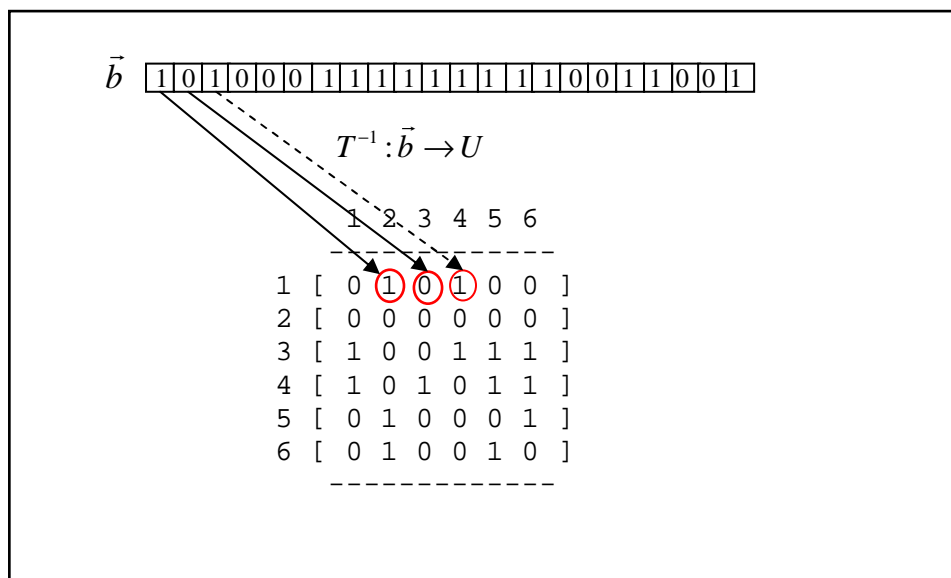
AG jest algorytmem iteracyjnym, w którym podczas każdej iteracji ocenie zostają poddane wszystkie chromosomy reprezentujące rozwiązania cząstkowe tak, jak zostało to przedstawione na rysunku poniżej.



Rys. 2. Schemat blokowy działania algorytmu genetycznego

Ze zbioru chromosomów, które uzyskały najwyższe noty, losowana jest tzw. *pula rodzicielska*, na podstawie której tworzona jest kolejna generacja chromosomów. Zatem podstawowym założeniem jest to, że oceny poszczególnych rozwiązań cząstkowych są rzetelne i obrazują, które z nich wskazują właściwy kierunek poprawy. Ocena chromosomów (wektorów bitowych) może być

realizowana bez wykonywania przekształcenia odwrotnego, tj. $T^{-1} : \vec{b} \rightarrow U$, gdyż odwzorowanie struktur OD do wektorów bitowych jest jednoznaczne. W celach poglądowych wspomniane przekształcenie odwrotne zostało zaprezentowane na przykładzie na rysunku 3.



Rys. 3. Przekształcenie wektora bitowego do zbioru łuków

Ze względu na fakt, że ocena poszczególnych chromosomów jest najczęściej wykonywaną operacją w czasie działania algorytmu genetycznego, stanowi ona o jego efektywności, a co za tym idzie przydatności przy rozwiązywaniu realnych problemów. W dalszej części artykułu zostały przedstawione dwie funkcje przystosowania zdefiniowane odpowiednio dla struktur opiniowania diagnostycznego PMC i BGM.

5. Konstrukcja funkcji przystosowania

W przypadku struktur opiniowania diagnostycznego typu PMC funkcja przystosowania została wyznaczona w oparciu o warunek wystarczający (2) w taki sposób, by nadawać wyższe noty rozwiązaniom cząstkowym posiadającym wystarczającą liczbę niepowtarzalnych poprzedników. Postać opracowanej funkcji jest następująca:

$$f(G) = \sum_{p=0}^{m-1} \sum_{E' \subset Z_p} k(E'),$$

gdzie:

E' – podzbiór zbioru E , przy czym struktura OD to $G = \langle E, U \rangle$,

Z_p – zbiór węzłów taki, że $\forall E' \subset Z_p : |E'| = |E| - 2 \cdot m + p$,

m – parametr określający krotność diagnozowalności struktury OD,

$k(E')$ – funkcja zdefiniowana następująco:

$$k(E') = \begin{cases} 1 & \text{dla } \Gamma(E') = p \\ -|p - \Gamma(E')| & \text{dla } p \neq \Gamma(E'). \end{cases}$$

Zdefiniowana w powyższy sposób funkcja przystosowania gwarantuje, iż struktury odnajdywane przez AG będą m -diagnozowalne, jednakże nie daje ona żadnej kontroli nad ich pozostałymi własnościami. Zgodnie z [4] najbardziej pożądane są struktury optymalne, gdyż są one również strukturami ekonomicznymi, tj. strukturami o najniższym koszcie połączeń. W związku z powyższym rozsądne wydaje się zmodyfikowanie zaprezentowanej funkcji przystosowania w taki sposób, by najwyższe oceny uzyskiwały optymalne w sensie diagnostyki struktury OD. Dokonać tego można poprzez dodanie do funkcji przystosowania członu podwyższającego ocenę w przypadku, gdy badane rozwiązanie wykazuje własności typowe dla struktur optymalnych. Na podstawie [5] należy przyjąć, iż cechą charakteryzującą pożądane rozwiązania (optymalne) jest $\forall e \in E : \mu^-(e) = \mu^+(e) = m$, zatem funkcja przystosowania może zostać zmodyfikowana do następującej postaci:

$$f(G) = \sum_{e \in E} g(e) + \sum_{p=0}^{m-1} \sum_{E' \subset Z_p} k(E'),$$

gdzie: funkcja $g_i(e)$ jest zdefiniowana następująco:

$$g(e) = \begin{cases} 1 & \text{dla } \mu^-(e) = \mu^+(e) = m \\ -|2m - (\mu^-(e) + \mu^+(e))| - |\mu^-(e) - \mu^+(e)| & \text{dla } \mu^-(e) \neq m, \end{cases}$$

a $\mu^+(e)$ oznacza stopień wyjściowy węzła.

Dla przypadku, gdy rozpatrywany jest jedynie zbiór struktur opiniowania diagnostycznego typu BGM, które nie są jednocześnie strukturami typu PMC, funkcja przystosowania powinna być skonstruowana na podstawie wcześniej wymienionej zależności, z której jednoznacznie wynika, że dowolne dwa węzły

grafu G nie mogą mieć takiego samego zbioru poprzedników. Proponowana funkcja przystosowania ma następującą postać:

$$f(G) = \sum_{\langle e, e' \rangle \in U} w(\Gamma^-(e), \Gamma^-(e')),$$

gdzie: $\Gamma^-(e)$ – zbiór poprzedników wężła e , a $w(\Gamma^-(e), \Gamma^-(e'))$ oznacza funkcję zdefiniowaną w następujący sposób:

$$w(\Gamma^-(e), \Gamma^-(e')) = \begin{cases} 1 & \text{dla } \Gamma^-(e) \neq \Gamma^-(e') \\ 0 & \text{w p.p.} \end{cases}$$

Podobnie jak przy strukturach typu PMC, także i w tym przypadku podstawowa postać funkcji przystosowania przyznaje pozytywne oceny wszystkim m -diagnozowalnym strukturom, jednakże bez uwzględnienia ich dodatkowych własności. W celu zwiększenia efektywności wyznaczania przez AG struktury optymalnej typu BGM zgodnej z zadaną topologią konieczne jest dodanie do funkcji przystosowania dodatkowego członu w postaci wcześniej zdefiniowanej funkcji $g_i(e)$. W takim przypadku funkcja przystosowania przyjmie następującą postać:

$$f(G) = \sum_{e \in E} g(e) + \sum_{\langle e, e' \rangle \in U} w(\Gamma^-(e), \Gamma^-(e')).$$

Warto zwrócić uwagę na fakt, że funkcja przystosowania powinna jednoznacznie określać lepszy chromosom poprzez przyznanie mu wyższej noty. W zaprezentowanym rozwiązaniu wątpliwość może budzić operacja dodawania, która może doprowadzić do sytuacji, w której dwie struktury o różnych własnościach zostaną identycznie ocenione. Jest to jednakże działanie zamierzone, gdyż badane struktury mogą w równym stopniu spełniać różne wcześniej wysunięte postulaty, a przez to mogą być w równym stopniu (w znaczeniu diagnostyki) przydatne w dalszej ewolucji. Badania empiryczne potwierdziły słuszność poczynionego założenia.

6. Uwzględnienie uogólnionych kosztów testów

Przedstawione w poprzednich rozdziałach funkcje przystosowania, umożliwiające odnajdywanie m -diagnozowalnych SOD typów PMC oraz BGM, opierają się na założeniu, że wszystkie testy posiadają jednakowy koszt. W celu uwzględnienia w ostatecznym rozwiązaniu dodatkowych warunków, określonych w postaci uogólnionego kosztu poszczególnych testów, należy rozwinąć

wcześniej przedstawione funkcje przystosowania do poniżej przedstawionych postaci:

- dla struktur typu PMC:

$$f(G) = \sum_{p=0}^{m-1} \sum_{E' \subset Z_p} k(E') - \sum_{i=1}^{|E|} \sum_{u \in U_i} n(r(u))$$

lub dla struktur optymalnych

$$f(G) = \sum_{e \in E} g(e) + \sum_{p=0}^{m-1} \sum_{E' \subset Z_p} k(E') - \sum_{i=1}^{|E|} \sum_{u \in U_i} n(r(u))$$

- dla struktur typu BGM:

$$f(G) = \sum_{\langle e, e' \rangle \in U} w(\Gamma^-(e), \Gamma^-(e')) - \sum_{i=1}^{|E|} \sum_{u \in U_i} n(r(u))$$

lub dla struktur optymalnych

$$f(G) = \sum_{e \in E} g(e) + \sum_{\langle e, e' \rangle \in U} w(\Gamma^-(e), \Gamma^-(e')) - \sum_{i=1}^{|E|} \sum_{u \in U_i} n(r(u)).$$

W powyższych formułach U_i jest zbiorem testów realizowanych przez i -ty węzeł ($U_i \subset U$), a $r(u)$ – uogólnionym kosztem testu u . Dodatkowy człon postaci: $\sum_{u \in U_i} n(r(u))$ reprezentuje sumaryczne znormalizowane koszty wszystkich

łuków wychodzących z węzła i .

Warto zwrócić uwagę na fakt, że uogólniony koszt przypisany do linii transmisyjnych w prezentowanym rozwiązaniu posiada inną interpretację, niż w przypadku omawianym w [8]. Ze względu na konstrukcję chromosomów jednoznacznie opisującą strukturę logiczną systemu uogólnione koszty oznaczają jedynie ekonomiczny koszt wykorzystania poszczególnych połączeń (nie pełnią dodatkowej roli opisu topologii systemu).

7. Ocena przydatności algorytmu

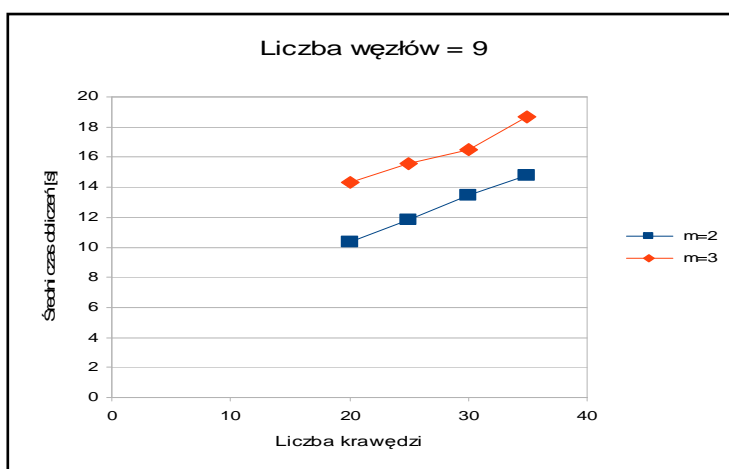
Przedstawiona wyżej adaptacyjna metoda projektowania struktur OD dla sieci teleinformatycznych o zadanej strukturze komunikacyjnej została zaimplementowana w języku C. W celu weryfikacji jej przydatności przeprowadzono szereg testów, których celem była ocena wpływu liczby węzłów i krawędzi (przy ustalonym parametrze m) grafu początkowego (reprezentującego sieć, dla której projektowana jest struktura OD) na czas

wyznaczenia struktury. Wszystkie testy zostały wykonane na komputerze klasy Pentium z zegarem 2 GHz przy wykorzystaniu parametrów AG przedstawionych w tabeli 1.

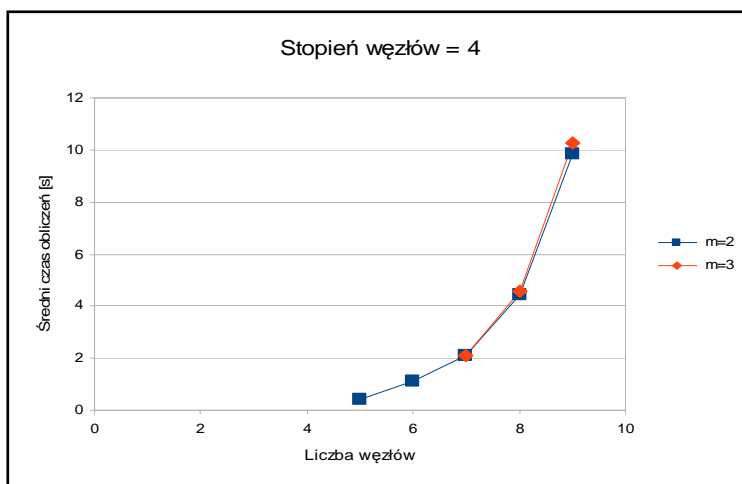
Tab. 1. Parametry algorytmu genetycznego

Parametr	Wartość
Populacja	100
Mutacja	0.01
Liczba przodków	3
Selekcja	ruletka
Krzyżowanie	dwupunktowe
Prawdopodobieństwo krzyżowania	0.95

Uzyskane wyniki przeprowadzonych testów jednoznacznie wskazują, że opracowana metoda jest liniowo zależna od liczby krawędzi (rys. 4) oraz wykładniczo od liczby węzłów (rys. 5). Warto zwrócić uwagę na fakt, że liczba węzłów powiązana jest z minimalną liczbą krawędzi występujących w grafie G poprzez własności (1), (2) dla struktur PMC i własności (3), (4) dla struktur BGM. Wspomniane własności przesądzają, że dla struktury t -diagnozowalnej stopnie wejściowe wszystkich węzłów muszą być większe lub równe t , a w przypadku struktury o najmniejszym możliwym uogólnionym koszcie powinno zachodzić $\forall e \in E : \mu^-(e) = \mu^+(e) = t \rightarrow \forall e \in E : \mu(e) = 2t$.



Rys. 4. Zobrazowanie wpływu liczby krawędzi na średni czas działania algorytmu



Rys. 5. Zobrazowanie wpływu liczby węzłów na średni czas działania algorytmu

Zauważmy, że eksperymenty zostały przeprowadzone dla struktur, w których liczba węzłów wynosiła od 5 do 9. W związku z powyższym przestrzeń poszukiwań rozwiązania wynosiła odpowiednio od 2^{20} do 2^{72} . Zależność ta wynika bezpośrednio z maksymalnej liczby kombinacji opiniowań, które mogą być określone w macierzy przejść $M'_{y \times z=(y-1)}$ (diagonala macierzy $M'_{y \times z=(y-1)}$ z założenia zawiera same zera). Dokładne przełożenie liczby węzłów w projektowanej strukturze (od 5 do 9 węzłów) na wielkość przestrzeni poszukiwań rozwiązania zostało przedstawione w tabeli 2.

Tab. 2. Zależność liczby możliwych rozwiązań od liczby węzłów

Liczba węzłów	5	6	7	8	9
Liczba możliwych rozwiązań	2^{20}	2^{30}	2^{42}	2^{56}	2^{72}

Badania przeprowadzono dla przypadków skrajnych, gdy możliwe było wyznaczenie tylko kilku m -diagnostowalnych struktur OD o zadanej topologii. W sytuacjach zazwyczaj spotykanych liczba możliwości połączeń między węzłami jest znacznie większa, a w związku z tym efektywność proponowanej metody wzrasta.

8. Podsumowanie

Przedstawiona w artykule metoda pozwala na wyznaczenie najtańszych m -diagnostowalnych struktur opiniowania diagnostycznego z uwzględnieniem istniejącej topologii sieci teleinformatycznej. Może ona znaleźć zastosowanie w praktyce, gdy zachodzi potrzeba wyznaczenia nowej struktury diagnostycznej sieci po dokonaniu jej rekonfiguracji np. wskutek wykrycia niezdatności określonych węzłów.

Wyniki przeprowadzonych eksperymentów wskazują, iż opracowany algorytm sprawdza się zarówno w przypadku struktur regularnych, jak i nieregularnych.

Zastosowanie zaproponowanych heurystyk daje zadowalające rezultaty przy dopuszczalnych (w praktyce eksploatacji sieci teleinformatycznych) czasach poszukiwania rozwiązania.

Literatura

- [1] ARCIUCH A., *Wyznaczanie 3-optimalnych struktur opiniowania diagnostycznego typu PMC*, Biuletyn IAIr, nr 22, Warszawa, 2005, str. 33-54.
- [2] ARCIUCH A., *Metody projektowania 3-optimalnych struktur opiniowania diagnostycznego, spełniających określone wymagania*, WAT, Warszawa, 2007.
- [3] BARSİ F., GRANDONI P., MAESTRINI A., *Theory of Diagnosability of Digital Systems*, IEEE Transactions on Computers, 1976, C-25, 6.
- [4] KULESZA R., *Problemy przeliczania optimalnych struktur opiniowania diagnostycznego*, Biuletyn IAIr, nr 20, Warszawa, 2004, str. 3-21.
- [5] KULESZA R., *Podstawy diagnostyki sieci logicznych i komputerowych*, IAIr, Wydział Cybernetyki WAT, Warszawa, 2000.
- [6] GOLDBERG D.E., *Algorytmy genetyczne i ich zastosowania*, WNT, Warszawa 1995, 1998, 2003.
- [7] STRZELECKI Ł., RENCZEWSKI K., *Adaptacyjna metoda wyznaczania ekonomicznych m -diagnostowalnych struktur opiniowania diagnostycznego typu PMC*, Biuletyn IAIr, nr 25, Warszawa, 2008, str. 139-153.
- [8] STRZELECKI Ł., ZIELIŃSKI Z., *Wyznaczanie najtańszych struktur opiniowania diagnostycznego metodą adaptacyjną*, [w:] Systemy Czasu Rzeczywistego. Postępy badań i zastosowania (red. Zieliński Z.), WKŁ, 2009, str. 227-241.
- [9] PREPARATA F.P., METZE G., CHIEN R.T., *On the Connection Assignment Problem of Diagnosable Systems*, IEEE Transactions on Computers, 1967, 6.

- [10] HAKIMI S.L., AMIN A.T., *Characterisation of Connection Assignment Problem of Diagnosable Systems*, IEEE Transactions on Computers, 1974, 1.
- [11] ZIELIŃSKI Z., *Komputerowo wspomagane wyznaczanie najtańszych 2-diagnozowalnych struktur typu PMC*, Biuletyn IAIr, nr 18, Warszawa, 2003, str. 33-44.

Diagnostics structures designing with the use of genetics algorithm

ABSTRACT: The paper deals with the problem of designing diagnostic structures for computer networks with a fixed topology. An adaptive method for determining the cheapest structure using genetic algorithm is presented.

KEYWORDS: system level diagnosis, diagnostic structures, PMC and BGM models, genetics algorithm

Praca naukowa finansowana ze środków przeznaczonych na naukę w latach 2008-2011 jako projekt badawczy nr O 514051135.

Praca wpłynęła do redakcji: 02.11.2009.