BIULETYN INSTYTUTU AUTOMATYKI I ROBOTYKI
NR 18, 2003

Ocena wpływu procesu testowania na niezawodność modułu programowego

Kazimierz WORWA

Zakład Inżynierii Oprogramowania, Instytut Systemów Informatycznych,
Wojskowa Akademia Techniczna, ul. Kaliskiego 2, 00-908 Warszawa

STRESZCZENIE: W artykule przestawiono propozycję modelu wzrostu niezawodności modułu w procesie jego testowania. W celu uwzględnienia, bardzo często występującego w praktyce testowania oprogramowania, zjawiska wykrywania przez różne zestawy danych testowych tych samych błędów wprowadzono pojęcie tzw. macierzy charakterystycznej testowanego modułu. Na podstawie analizy zależności macierzy charakterystycznej i przyrostu wartości wskaźnika niezawodności modułu przedstawiono praktyczne oszacowania tego przyrostu dla wybranych postaci macierzy charakterystycznej testowanego modułu.

1. Wprowadzenie

Zasadniczym celem testowania oprogramowania jest wykrycie i usunięcie jak największej liczby błędów popełnionych w trakcie realizacji wcześniejszych etapów procesu jego wytwarzania. Wykrycie i usunięcie określonej liczby błędów przyczynia się do zwiększenia niezawodności oprogramowania. Liczba wykrytych błędów silnie zależy od zakresu, dokładności i organizacji prac związanych z testowaniem. Praktyka produkcji oprogramowania pokazuje, że prace te są bardzo czasochłonne i wymagają wysokich nakładów finansowych. Powoduje to m.in., że etap testowania charakteryzuje się bardzo znacznym udziałem w ogólnym koszcie wytworzenia modułu. W pracach [7,12] wskazuje się, że w przypadku dużego i złożonego systemu programowego koszt testowania może stanowić 40-70% łącznego kosztu jego produkcji. Fakt ten jest bezpośrednią konsekwencją dużej czasochłonności prac prowadzonych na tym etapie, konieczności zaangażowania doświadczonych i wykwalifikowanych programistów oraz dużego zużycia czasu pracy wykorzystywanego zestawu komputerowego.

Z uwagi na wspomniany duży udział kosztu procesu testowania w ogólnym koszcie produkcji oprogramowania, realizacja etapu testowania powinna być poprzedzona jego właściwym zaplanowaniem i przygotowaniem. Planowanie procesu testowania wymaga w szczególności określenia organizacji prac, składających się na etap testowania oraz sposobu generowania wykorzystywanego zbioru zestawów danych testowych. Podejmowane podczas planowania procesu testowania decyzje mają zasadniczy wpływ na czas i koszt tego procesu. Dla zwiększenia poprawności i trafności tych decyzji pomocne jest wykorzystywanie metod formalnych, w tym zwłaszcza metod badań operacyjnych i analizy systemowej.

Podstawowym celem rozważań zawartych w artykule jest ocena przewidywanego wpływu przyjętej strategii testowania modułu na poziom jego niezawodności. Ocena ta jest prowadzona dla potrzeb wspomagania planowania realizacji procesu testowania. Jej wyniki mogą umożliwić w szczególności określenie nakładów czasowo-finansowych, wymaganych dla realizacji testowania według określonej strategii. Celowość przeprowadzenia takich badań wynika z faktu, że pomimo niekwestionowanych potrzeb w tym zakresie, w praktyce brak jest modelu procesu testowania oprogramowania, umożliwiającego efektywne wspomaganie planowania realizacji procesu testowania. Znane i wykorzystywane w praktyce testowania modele niezawodności, umożliwiające ocenę poziomu niezawodności oprogramowania w trakcie realizacji testowania lub po jego zakończeniu, są nieprzydatne lub nie mogą być wykorzystywane na etapie planowania procesu testowania. Sytuacja ta wynika z dwóch zasadniczych powodów:

– dla racjonalnego planowania procesu testowania zależności analityczne, określające wykorzystywane w ramach tych modeli charakterystyki niezawodnościowe testowanego oprogramowania, nie są jawnymi funkcjami wielkości opisujących sposób realizacji procesu testowania, takich jak czas testowania czy liczebność zbioru testowych;
– istniejące modele niezawodności umożliwiają określenie, zasadniczych dla racjonalnego planowania procesu testowania oprogramowania, charakterystyk takich jak: przewidywany czas testowania, liczebność wymaganego zbioru testowych, oczekiwana liczba wykrytych błędów itp., dopiero wówczas, gdy oprogramowanie to osiągnie poziom komputerowej wykonywalności [1,2,4,8,9,10,12,15,21].

W artykule przedstawiony jest matematyczny model wzrostu niezawodności modułu w procesie jego testowania, przy czym proponowany wskaźnik niezawodności modułu oparty został na prawdopodobieństwie poprawnego wykonania się modułu dla pojedynczego zestawu danych wejściowych (*run reliability*). Dla potrzeb konstrukcji modelu przyjęto ważne dla dalszej części prezentowanych rozważań określenia strategii testowania i macierzy charakterystycznej testowanego modułu. W dalszej części pracy przeanalizowany został wpływ macierzy charakterystycznej testowanego modułu na wartość wskaźnika jego niezawodności dla danej strategii testowania. Przedstawiona została także ocena wpływu strategii testowania na wartość wskaźnika niezawodności testowanego modułu, przy czym ocena ta obejmuje przypadek braku możliwości wykrywania przez różne zestawy danych tych samych błędów oraz przypadek, gdy taka sytuacja jest możliwa. Przypadki te opisywane są odpowiednio dobranymi macierzami charakterystycznymi rozpatrywanego modułu. Rozpatrzony został także, jak się wydaje najbliższy realiom praktyki testowania oprogramowania, przypadek tzw. dwumianowej macierzy charakterystycznej testowanego modułu.

2. Schemat procesu testowania modułu

Przebieg prac składających się na etap testowania modułu zależy od przyjętego schematu procesu testowania. Schemat ten określa w szczególności organizację procesu testowania oraz procedurę jego realizacji.

W dalszych rozważaniach przyjmuje się, że schemat procesu testowania modułu oparty jest na sekwencji tzw. cykli testowania, każdy z których obejmuje:

1) wykonanie (uruchomienie) modułu dla pewnej liczby uprzednio przygotowanych zestawów danych testowych,
2) ocenę uzyskanych rezultatów oraz lokalizację i usunięcie ewentualnie wykrytych błędów.

Zgodnie z przyjętym schematem testowania pierwszy etap każdego cyklu procesu testowania polega na uruchomieniu modułu dla pewnej liczby przygotowanych zestawów danych testowych, przy czym do momentu zakończenia wykonywania modułu dla wszystkich, przewidzianych w danym cyklu, zestawów danych kod źródłowy testowanego modułu pozostaje nie zmieniony (ewentualnie wykrywane błędy są jedynie rejestrowane). Ingerencja w kod źródłowy testowanego modułu, związana z koniecznością usunięcia wykrytych błędów, następuje w ramach drugiego etapu cyklu testowania. Istotą tego etapu cyklu procesu testowania modułu jest ocena rezultatów testowania, przeprowadzonego w ramach pierwszego etapu, oraz lokalizacja i usunięcie ewentualnie wykrytych błędów.

Przyjęcie opisanego schematu testowania oznacza, że w ramach określonego cyklu procesu testowania warunki wykonywania modułu dla poszczególnych zestawów danych testowych są dokładnie takie same, natomiast po przejściu do kolejnego cyklu moduł (a ściślej jego kod źródłowy) może być zmieniony (w wyniku usunięcia błędów wykrytych w poprzednim cyklu).

Warto podkreślić, że w zasadzie każdy, realizowany w praktyce proces testowania oprogramowania, przebiega zgodnie z wyżej przedstawionym schematem testowania. Ewentualne różnice wynikają z liczby zestawów danych, wykorzystywanych w pierwszym etapie każdego cyklu. Ogólnie liczba tych zestawów może być różna w każdym cyklu. Często przyjmuje się jednak, że w każdym cyklu wykorzystuje się taką samą liczbę zestawów danych testowych [20].

Warto nadmienić, że użytkowa eksploatacja oprogramowania systemów informatycznych, która może być traktowana jako kolejny specyficzny cykl procesu testowania, także przebiega zgodnie z wcześniej opisanym, ogólnym schematem testowania. Oprogramowanie takie jest używane często przez wielu niezależnych użytkowników. W przypadku wykrycia przez któregoś z nich błędu, powiadamiany jest producent oprogramowania, który dokonuje rejestracji tego błędu. Po pewnym czasie następuje usunięcie wszystkich, wykrytych w ten sposób błędów, a wolne od nich oprogramowanie jest dostarczane lub oferowane użytkownikom jako jego kolejna wersja.

3. Macierz charakterystyczna modułu

Z opisu przyjętego schematu procesu testowania modułu wynika, że w ramach pojedynczego cyklu testowania lokalizacja i usuwanie ewentualnie wykrytych błędów następuje dopiero po uruchomieniu modułu dla wszystkich, zaplanowanych dla danego cyklu, zestawów danych. Przyjęcie opisywanego schematu oznacza w praktyce, że w ramach określonego cyklu kolejny test wykonywany jest bez analizy i oceny poprawności wykonania się poprzedniego testu. W opisywanej sytuacji możliwe są zatem takie przypadki, w których określony błąd programowy zostanie wykryty przez kilka (tj. dwa lub więcej) różnych, wykorzystywanych w ramach tego samego cyklu, zestawów danych. Opisywane zjawisko może w oczywisty sposób zmniejszać efektywność procesu testowania, ponieważ te zestawy danych, spośród wykorzystywanych w danym cyklu, które wskazują jedynie na błędy wykryte przez inne wcześniejsze zestawy danych są nieefektywne. Warto zauważyć, że wspomnianej, możliwej nieefektywności procesu testowania w ramach założonego schematu, można uniknąć, przyjmując, że analiza wyników testowania i usuwanie ewentualnie wykrytych błędów odbywać się będzie każdorazowo po wykonaniu pojedynczego testu. Testowanie takie jest jednakże rzadko realizowane w praktyce, z uwagi na jego nieefektywność ekonomiczną, wyrażającą się w bardzo znaczącym wzroście wymaganych nakładów czasowo-finansowych.

Wydłużenie czasu procesu testowania wynika z faktu, że po każdym wykrytym błędzie proces ten jest wstrzymywany do czasu jego zlokalizowania i usunięcia. Z kolei wzrost kosztu testowania jest bezpośrednią konsekwencją wzrostu kosztów kompilacji i powtórnego testowania, wymaganego dla oceny poprawności wprowadzonych zmian w kodzie źródłowym testowanego modułu.

Wspomniana możliwość wykrywania przez różne zestawy danych, wykonywane w ramach jednego cyklu, tych samych błędów oznacza, że liczba zestawów danych, dla których otrzymano niepoprawne, tj. inne od oczekiwanych, wyniki, może być większa od łącznej liczby różnych błędów, wykrytych w danym cyklu testowania. Można ogólnie stwierdzić, że możliwość wykrywania przez różne zestawy danych, wykorzystywane w ramach określonego cyklu procesu testowania, tych samych błędów zależy m.in. od:

– struktury logicznej testowanego modułu, określonej przez liczbę możliwych dróg, które może uaktywnić pojedynczy zestaw danych testowych,
– sposobu tworzenia (projektowania) wykorzystywanego w procesie testowania zbioru danych testowych.

Ocenia się [7,12], że im większy jest zbiór instrukcji testowanego modułu, stanowiący część wspólną poszczególnych dróg, tym większe jest prawdopodobieństwo zdarzenia polegającego na tym, że kilka różnych zestawów danych, spośród wykorzystywanych w ramach tego samego cyklu procesu testowania modułu, wykryje ten sam błąd.

Należy podkreślić, że pomimo iż możliwość wielokrotnego wykrywania tych samych błędów przez różne zestawy danych testowych, wykorzystywanych w ramach określonego cyklu, może w istotny sposób wpływać na efektywność całego procesu testowania, większość modeli wzrostu niezawodności oprogramowania w procesie testowania nie uwzględnia tego zjawiska. Dotyczy to w szczególności modeli niezawodności oprogramowania typu dwumianowego [7,16,21]. Główną przyczyną tego stanu rzeczy są trudności formalnego opisu i analizy tego zjawiska. Jedną z pierwszych prób jego uwzględnienia podjęto w pracy [20]. W cytowanej pracy rozpatrywany jest schemat, zakładający realizację pewnej liczby cykli testowania (*testing rounds*), z wykorzystaniem wielu zestawów danych testowych (*testing cases*) w ramach każdego z nich, przy czym dla uproszczenia założono, że w każdym cyklu wykorzystuje się taką samą liczbę zestawów danych. Uwzględnia się przy tym możliwość wielokrotnego wykrycia tego samego błędu przez różne, wykonywane w ramach danego cyklu, zestawy danych testowych. Testowanie to określane jest w cytowanej pracy terminem *recapture testing*.

Niech $P=\left[p_{nm}\right],\ 0\le p_{nm}\le 1,\ n,m\in\{0,1,2,...\}$, oznacza macierz nieskończoną, element p_{nm} której oznacza prawdopodobieństwo zdarzenia polegającego na tym, że w jednym cyklu procesu testowania wykrytych zostanie

n nowych błędów, pod warunkiem, że w ramach tego cyklu dla m zestawów danych otrzymane zostaną niepoprawne wyniki.

Zakładając, że w przypadku każdego testu, który dał niepoprawny wynik, nastąpi wykrycie powodującego taki stan błędu oraz że każdy test może wykryć co najwyżej jeden błąd, można napisać:

$$
\begin{aligned}
&p_{00} = p_{11} = 1 \\
&p_{nm} = 0 \quad dla\ n=0 \text{ i } m>0 \text{ lub dla } n > m, \\
&0 < p_{nm} < 1 \quad \text{dla } n \leq m,\ n > 0\,.
\end{aligned}
\tag{1}
$$

Uwzględniając własności elementów p_{nm}, opisane zależnością (1), macierz P można przedstawić w następującej potaci:

$$
P = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & \dots \\
0 & 1 & p_{12} & p_{13} & p_{14} & \dots \\
0 & 0 & p_{22} & p_{23} & p_{24} & \dots \\
0 & 0 & 0 & p_{33} & p_{34} & .. \\
0 & 0 & 0 & 0 & p_{44} & \dots \\
 & & \dots & & &
\end{bmatrix}.
\tag{2}
$$

Wielkości p_{nm} spełniają warunek:

$$
\sum_{n=0}^{\infty} p_{nm} = 1, \qquad m \in \{0,1,2,\dots\}.
\tag{3}
$$

Warunek (3) oznacza, że suma elementów w każdej kolumnie macierzy P jest równa 1.

Dla ustalonej metody projektowania zestawów danych testowych wartości prawdopodobieństw p_{nm}, tworzących macierz P, zależne są od struktury logicznej kodu źródłowego testowanego modułu i od poziomu jego niezawodności. Można ogólnie stwierdzić, że prawdopodobieństwa p_{nm} zależą przede wszystkim od:

- liczby dróg logicznych, łączących wierzchołek wejściowy z wierzchołkiem wyjściowym,
- długości tych dróg, mierzonej np. liczbą instrukcji modułu wykonywanych w przypadku ich uaktywnienia,
- stopnia wzajemnego pokrywania się poszczególnych dróg logicznych, miarą którego jest liczba instrukcji modułu wchodzących w skład dwóch lub większej liczby dróg.

Przykładowo, jeżeli w testowanym module każda instrukcja (z wyjątkiem instrukcji tworzących wierzchołki wejściowy i wyjściowy) należy tylko do jednej drogi logicznej, tzn. że testowany moduł składa się ze zbioru dróg rozłącznych, wówczas należy oczekiwać, że wartości prawdopodobieństw p_{nm}, $n \leq m$, będą skupione na głównej przekątnej macierzy *P* lub w jej pobliżu (nad główną przekątną). Z kolei, jeśli w testowanym module poszczególne drogi nakładają się, tzn. wiele dróg „przechodzi" przez te same instrukcje lub ich grupy, wówczas należy oczekiwać, że wartości prawdopodobieństw p_{nm}, $n \leq m$, będą skupione w początkowych wierszach macierzy *P* (z zachowaniem warunków (1)-(3).

Z wymienionych względów w dalszych rozważaniach macierz *P* nazywana będzie macierzą charakterystyczną testowanego modułu.

Określenie schematu testowania modułu wymaga podania liczby cykli oraz liczebności zbiorów danych testowych, wykorzystywanych w poszczególnych cyklach. Podane wielkości wyznaczają tzw. strategię testowania.

Niech *S* oznacza strategię testowania modułu określoną następująco:

$$S = (K, (L_1, L_2, ..., L_k, ..., L_K)), \tag{4}$$

gdzie:

K - liczba cykli procesu testowania modułu,

L_k - liczba zestawów danych, wykorzystywanych w *k*-tym cyklu procesu testowania modułu, $L_k > 0$.

Zgodnie z przyjętym schematem testowania modułu poprawianie błędów wykrytych w *k*-tym cyklu procesu testowania modułu odbywa się dopiero po uruchomieniu modułu dla wszystkich L_k zestawów danych, $k = \overline{1,K}$.

Niech ***S*** oznacza zbiór wszystkich możliwych strategii testowania modułu, tj.:

$$\boldsymbol{S} = \{S = (K,\ (L_1, L_2, ..., L_k, ..., L_K)) : \quad L_k > 0,\ \ k = \overline{1,K},\ \ K > 0\}. \tag{5}$$

Niech $M_k(S,P)$ oznacza liczbę zestawów danych, spośród przewidzianych do wykorzystania w ramach *k*-tego cyklu procesu testowania modułu o macierzy charakterystycznej *P*, realizowanego według strategii *S*, dla których wystąpią niepoprawne, tj. niezgodne z oczekiwanymi, wyniki.

Niech $N_k(S,P)$ oznacza liczbę różnych błędów, wykrycie których jest spodziewane w wyniku wykonania *k*-tego cyklu procesu testowania modułu, o macierzy charakterystycznej *P*, realizowanego według strategii *S*.

Ponieważ wielkość $M_k(S,P)$ może obejmować także zestawy danych, wykonanie których powoduje wykrycie tych samych błędów, zachodzi:

$$M_k(S,P) \geq N_k(S,P), \quad k=\overline{1,K}\,.$$

Wielkości $M_k(S,P)$, $N_k(S,P)$ są zmiennymi losowymi, przy czym ich rozkład łączny może być określony następująco:

$$\begin{aligned} &Pr\{\,N_k(\,S,P\,)=n_k\,,M_k(\,S,P\,)=m_k\,\}= \\ &\qquad Pr\{\,N_k(\,S,P\,)=n_k\,/M_k(\,S,P\,)=m_k\,\}\,Pr\{\,M_k(\,S,P\,)=m_k\,\}\,. \end{aligned} \tag{6}$$

Rozkład prawdopodobieństwa zmiennej losowej $N_k(S,P)$ może być określony jako rozkład brzegowy w rozkładzie dwuwymiarowej zmiennej losowej $(N_k(S,P),M_k(S,P))$:

$$\begin{aligned} &Pr\{\,N_k(\,S,P\,)=n_k\,\}= \\ &\quad =\sum_{m_k=0}^{L_k} Pr\{\,N_k(\,S,P\,)=n_k\,/M_k(\,S,P\,)=m_k\,\}\,Pr\{\,M_k(\,S,P\,)=m_k\,\}\,. \end{aligned} \tag{7}$$

Niech $N(S,P)$ oznacza ogólną liczbę różnych błędów, wykrycie których jest spodziewane w procesie testowania modułu o macierzy charakterystycznej P, realizowanego według strategii S. Zgodnie z wcześniejszymi uwagami, na etapie planowania procesu testowania modułu zasadne jest przyjęcie założenia o losowym charakterze wielkości $N(S,P)$. Uwzględniając bowiem przyjęte określenie strategii testowania (4) wielkość $N(S,P)$ można przedstawić następująco:

$$N(S,P)=\sum_{k=1}^{K} N_k(S,P)\,. \tag{8}$$

Uwzględniając przyjęty schemat testowania rozkład prawdopodobieństwa zmiennej losowej $N(S,P)$ można określić następująco:

$$\begin{aligned} &Pr\{N(S,P)=n\}= \\ &\quad =\sum_{n_1+n_2+\ldots+n_K=n} Pr\{N_1(S,P)=n_1,N_2(S,P)=n_2,\ldots,N_K(S,P)=n_K\,\}\,, \end{aligned} \tag{9}$$

$$n=0,1,2,\ldots$$

gdzie $Pr\{N_1(S,P)=n_1,N_2(S,P)=n_2,\ldots,N_K(S,P)=n_K\}$ określa łączny rozkład K-wymiarowej zmiennej losowej $(N_1(S,P),N_2(S,P),\ldots,N_K(S,P))$.

Rozkład $Pr\{N_1(S,P)=n_1\,,N_2(S,P)=n_2\,,\ldots,N_K(S,P)=n_K\,\}$ może być określony następująco:

$$
\begin{aligned}
&Pr\{\, N_1(S,P)=n_1, N_2(S,P)=n_2,\dots,N_K(S,P)=n_K \,\} = \\
&\quad = Pr\{N_K(S,P)=n_K \big/ N_i(S,P)=n_i, i=\overline{1,K-1}\}\cdot \\
&\qquad\qquad \cdot Pr\{N_1(S,P)=n_1,\dots,N_{K-1}(S,P)=n_{K-1}\} = \cdots \\
&\cdots = Pr\{N_1(S,P)=n_1\} \prod_{l=2}^{K} Pr\{N_l(S,P)=n_l \big/ N_i(S,P)=n_i, i=\overline{1,l-1}\}
\end{aligned}
$$

czyli

$$
\begin{aligned}
&\Pr\{N_1(S,P)=n_1, N_2(S,P)=n_2,\dots,N_K(S,P)=n_K\} = \\
&= \Pr\{N_1(S,P)=n_1\} \prod_{l=2}^{K} \Pr\{N_l(S,P)=n_l/N_i(S,P)=n_i, i=\overline{1,l-1}\}\,.
\end{aligned}
\tag{10}
$$

W celu uproszczenia zależności (10) przyjęte zostanie następujące założenie:

$$
Pr\{N_1(S,P)=n_1\} = Pr\{(N_1(S,P)=n_1 \big/ N_0(S,P)=n_0\}, \tag{11}
$$

gdzie, zgodnie z przyjętą konwencją, $N_0(S,P)$ oznacza liczbę błędów wykrytych przed rozpoczęciem testowania, $N_0(S,P)=0$.
Wykorzystując założenie (11), zależność (10) można przedstawić w postaci:

$$
\begin{aligned}
&Pr\{N_1(S,P)=n_1, N_2(S,P)=n_2,\dots,N_K(S,P)=n_K\} = \\
&\qquad = \prod_{l=1}^{K} Pr\{N_l(S,P)=n_l \big/ N_i(S,P)=n_i, i=\overline{1,l-1}\}.
\end{aligned}
\tag{12}
$$

Z zależności (7) wynika, że zachodzi:

$$
\begin{aligned}
&Pr\{N_k(S,P)=n_k \big/ N_i(S,P)=n_i, i=\overline{1,k-1}\} = \\
&= \sum_{m_k=0}^{L_k} Pr\{N_k(S,P)=n_k \big/ M_k(S,P)=m_k, N_i(S,P)=n_i, i=\overline{1,k-1}\}\cdot \\
&\qquad \cdot Pr\{M_k(S,P)=m_k \big/ N_i(S,P)=n_i, i=\overline{1,k-1}\},
\end{aligned}
\tag{13}
$$

przy czym, z uwagi na przyjęty schemat testowania, można napisać:

$$
\begin{aligned}
&Pr\{\, N_k(S,P)=n_k \big/ M_k(S,P)=m_k, N_i(S,P)=n_i, i=\overline{1,k-1}\} = \\
&\qquad = Pr\{N_k(S,P)=n_k \big/ M_k(S,P)=m_k\}.
\end{aligned}
\tag{14}
$$

Wykorzystując zależności (12), (7) oraz (13) i (14), otrzymuje się:

$$Pr\{N_1(S,P)=n_1, N_2(S,P)=n_2,...,N_K(S,P)=n_K\}=$$

$$=\prod_{k=1}^{K}\sum_{m_k=0}^{L_k} Pr\{N_k(S,P)=n_k \big/ M_k(S,P)=m_k\}\cdot \tag{15}$$

$$\cdot Pr\{M_k(S,P)=m_k \big/ N_i(S,P)=n_i, i=\overline{1,k-1}\}.$$

Ponieważ, zgodnie z określeniem elementów macierzy *P*, zachodzi:

$$Pr\{N_k(S,P)=n_k \big/ M_k(S,P)=m_k\}=p_{n_k m_k},\quad k=\overline{1,K} \tag{16}$$

więc zależność (15) można przedstawić w następującej postaci:

$$Pr\{N_1(S,P)=n_1, N_2(S,P)=n_2,...,N_K(S,P)=n_K\}=$$

$$=\prod_{k=1}^{K}\sum_{m_k=0}^{L_k} p_{n_k m_k}\, Pr\{M_k(S,P)=m_k \big/ N_i(S,P)=n_i, i=\overline{1,k-1}\}. \tag{17}$$

Dla ustalonej strategii testowania *S* i ustalonej wartości *k*, $k=\overline{1,K}$, prawdopodobieństwa warunkowe p_{n_k,m_k} tworzą następującą macierz $P_k(S)$, będącą podmacierzą macierzy charakterystycznej *P*, otrzymaną poprzez uwzględnienie jej pierwszych L_k+1 wierszy i kolumn, tj.:

$$P_k(S)=[p_{nm}]_{(L_k+1)\times(L_k+1)}\ , \tag{18}$$

gdzie L_k oznacza liczbę zestawów danych użytych w *k*-tym cyklu procesu testowania modułu, realizowanego według strategii *S*.
W rozpatrywanym przypadku macierze $P_k(S),\ \ k=\overline{1,K}$, są zatem macierzami postaci:

$$P_k(S)=\begin{bmatrix} 1 & 0 & 0 & 0 & ... & 0 \\ 0 & 1 & p_{12} & p_{13} & ... & p_{1L_K} \\ 0 & 0 & p_{22} & p_{23} & ... & p_{2L_K} \\ & & & & ... & \\ 0 & 0 & 0 & 0 & ... & p_{L_k L_k} \end{bmatrix}. \tag{19}$$

W wyniku wykonania etapu testowania, realizowanego według określonej strategii *S*, spodziewane jest wykrycie w testowanym module *N(S,P)* błędów, usunięcie których przyczyni się do wzrostu jego niezawodności. Liczba tych błędów, określona zależnością (8), zależy od poziomu niezawodności modułu w momencie rozpoczęcia procesu jego testowania oraz od przyjętej strategii testowania, w tym zwłaszcza od liczebności wykorzystywanego zbioru danych testowych.

4. Wskaźnik niezawodności modułu

Z praktyki testowania wynika, że wykonanie testowania zgodnie z określoną strategią nie daje pełnej gwarancji wykrycia wszystkich błędów, popełnionych na wcześniejszych etapach procesu wytwarzania modułu. Z uwagi na możliwość pozostania w module pewnej liczby błędów, wynik pojedynczego uruchomienia modułu o macierzy charakterystycznej *P*, po zakończeniu procesu testowania realizowanego według strategii *S*, można opisać następującą, binarną zmienną *X(S,P)*:

$$X(S,P) = \begin{cases} 1 & \text{jeśli pojedyncze wykonanie modułu o macierzy charakte-} \\ & \text{rystycznej } P\text{, testowanego według strategii } S\text{, jest} \\ & \text{poprawne,} \\ 0 & \text{w przeciwnym przypadku.} \end{cases}$$

Na etapie planowania procesu testowania modułu zasadne jest rozpatrywanie wielkości *N(S,P)* i *X(S,P)* - z uwagi na ich naturę - w kategoriach probabilistycznych, tj. jako zmiennych losowych. Zmienne losowe *X(S,P)*, *N(S,P)* nie są niezależne, gdyż - zgodnie z wcześniejszymi uwagami - wynik *X(S,P)* pojedynczego wykonania modułu, po zakończeniu procesu jego testowania realizowanego według strategii *S*, zależy od liczby błędów *N(S,P)*, wykrytych w trakcie tego procesu. Łączny rozkład prawdopodobieństwa dwuwymiarowej zmiennej losowej (*X(S,P)*, *N(S,P)*) można określić następująco:

$$Pr\{X(S,P) = x,\ N(S,P) = n\} = Pr\{\,X(S,P) = x \,/\, N(S,P) = n\,\}\ Pr\{N(S,P) = n\},$$

$$n \in \mathbf{N}(S,P),\ \ x \in \{0,1\}, \qquad (20)$$

gdzie ***N****(S,P)* oznacza zbiór wartości, które może przyjmować zmienna losowa *N(S,P)*:

$$\mathbf{N}(S,P) = \{\ 0,1,2,...,L(S)\ \}, \qquad (21)$$

przy czym

$$L(S) = \sum_{k=1}^{K} L_k\ . \qquad (22)$$

Wykorzystując zależność (20) można wyznaczyć rozkład prawdopodobieństwa zmiennej losowej *X(S,P)* jako rozkład brzegowy w rozkładzie dwuwymiarowej zmiennej losowej *(X(S,P), N(S,P))*:

$$\Pr\{X(S,P)=x\} = \\ = \sum_{n\in\mathbf{N}(S,P)} \Pr\{X(S,P)=x/N(S,P)=n\}\ \Pr\{N(S,P)=n\}\ ,\ \ x\in\{0,1\}. \tag{23}$$

W dalszych rozważaniach prawdopodobieństwo *Pr{X(S,P)=1}*, rozumiane jako prawdopodobieństwo poprawnego (tj. bezbłędnego), pojedynczego wykonania modułu o macierzy charakterystycznej *P*, po zakończeniu procesu jego testowania realizowanego według strategii *S*, nazywane będzie wskaźnikiem niezawodności modułu, oznaczanym przez *r(S,P)*. Zachodzi zatem:

$$r(S,P) = Pr\{X(S,P)=1\}. \tag{24}$$

Odpowiednio, prawdopodobieństwo warunkowe *Pr{X(S,P)=1/N(S,P)=n}*, rozumiane jako prawdopodobieństwo poprawnego, pojedynczego wykonania modułu o macierzy charakterystycznej *P*, pod warunkiem, że w procesie jego testowania, realizowanego według strategii *S*, wykryto i usunięto *n* błędów, nazywane będzie warunkowym wskaźnikiem niezawodności modułu, oznaczanym przez $r_n(S,P)$:

$$r_n(S,P) = Pr\{X(S,P) = 1/N(S,P) = n\}. \tag{25}$$

Zakładając, że każdy, wykorzystywany w procesie testowania, zestaw danych testowych (test) może wykryć co najwyżej jeden błąd oraz wykorzystując przyjęte określenie strategii testowania *S*, wskaźnik niezawodności testowanego modułu *r(S,P)* - zgodnie z (23) i (24) - może być określony w sposób następujący:

$$r(S,P) = \sum_{n=0}^{L(S)} r_n(S,P)\cdot Pr\{N(S,P)=n\}, \tag{26}$$

czyli, po uwzględnieniu zależności (9):

$$\mathrm{r(S,P)} = \sum_{\mathrm{n=0}}^{\mathrm{L(S)}} \mathrm{r_n(S,P)}\cdot \\ \sum_{\mathrm{n_1}+n_2+\ldots+n_K=n} \Pr\{\mathrm{N_1(S,P)=n_1,N_2(S,P)=n_2,\ldots,N_K(S,P)=n_K}\}\ , \tag{27}$$

gdzie wielkość *L(S)* oznacza maksymalną, dla przyjętej strategii testowania *S,* liczbę wykrytych błędów, określoną zależnością (22).

Uwzględniając zależność (17) otrzymuje się:

$$r(S,P) = \sum_{n=0}^{L(S)} r_n(S,P) \sum_{n_1+n_2+...+n_K=n} \prod_{k=1}^{K} \sum_{m_k=0}^{L_k} p_{n_k m_k} \Pr\{M_k(S,P) = m_k / N_i(S,P) = n_i,\ i = \overline{1,k\text{-}1}\}. \quad 28)$$

Usunięcie z testowanego modułu pewnej liczby błędów powoduje określony przyrost jego niezawodności. Oznacza to, że zachodzi:

$$r_n(S,P) = r + \Delta r(S,P,n)\,, \quad (29)$$

gdzie:

- r – wskaźnik niezawodności modułu o macierzy charakterystycznej P w momencie rozpoczęcia procesu jego testowania,
- $\Delta r(S,P,n)$ – przyrost wskaźnika niezawodności modułu o macierzy charakterystycznej P, będący wynikiem wykrycia i usunięcia $N(S,P) = n$ błędów.

Wyniki badań, przedstawione w pracach [3,5,11,12,14] pokazują, że przyrost $\Delta r(S,P,n)$ można przedstawić w postaci:

$$\Delta r(S,P,n) = (1 - r)(1 - e^{-\alpha n})\,, \quad (30)$$

gdzie wielkość α jest parametrem charakteryzującym "podatność" testowanego modułu na wzrost niezawodności (zależną m.in. od jego złożoności strukturalnej).

Po podstawieniu wyrażenia na $\Delta r(S,P,n)$ do zależności (29) otrzymuje się:

$$r_n(S,P) = 1 - (1 - r)e^{-\alpha n}\,. \quad (31)$$

Z zależności (28) wynika, że wskaźnik niezawodności modułu o macierzy charakterystycznej P, po zakończeniu procesu jego testowania realizowanego według strategii S, zależy od rozkładów warunkowych $r_n(S,P)$ oraz

$$Pr\{M_k(S,P) = m_k / N_i(S,P) = n_i, i = \overline{1,k - 1}\}\,.$$

Przyjęty schemat testowania, w którym prawdopodobieństwo poprawnego wykonania modułu dla pojedynczego zestawu danych testowych pozostaje stałe w ramach cyklu testowania, odpowiada znanemu w probabilistyce schematowi Bernoulliego. Prawdopodobieństwa warunkowe z wyrażenia (15) określone są zatem następującym rozkładem dwumianowym:

$$Pr\{M_k(S,P) = m_k / N_i(S,P) = n_i, i = \overline{1,k-1}\} =$$

$$= \binom{L_k}{m_k} [e^{-\alpha \sum_{i=1}^{k-1} n_i} (1-r)]^{m_k} [1-e^{-\alpha \sum_{i=1}^{k-1} n_i} (1-r)]^{L_k - m_k}, \quad (32)$$

$$m_k \in \{0,1,2,...,L_k\},\ k = \overline{1,K}.$$

Zatem:

$$Pr\{N_1(S,P) = n_1, N_2(S,P) = n_2, ..., N_K(S,P) = n_K\} =$$

$$= \prod_{k=1}^{K} \sum_{m_k=0}^{L_k} p_{n_k m_k} A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k), \quad (33)$$

gdzie

$$A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k) = \binom{L_k}{m_k} [e^{-\alpha \sum_{i=1}^{k-1} n_i} (1-r)]^{m_k} [1-e^{-\alpha \sum_{i=1}^{k-1} n_i} (1-r)]^{L_k - m_k}, \quad (34)$$

$$m_k \in \{0,1,2,...L_k\},\ k = \overline{1,K}.$$

Wielkość $A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k)$, $m_k \in \{0,1,2,...,L_k\}$, $n_i \in \{0,1,2,...,L_k\}$, $k \in \{1,2,...,K\}$, oznacza prawdopodobieństwo zdarzenia polegającego na tym, że spośród L_k zestawów danych, użytych w *k*-tym cyklu procesu testowania modułu, dla m_k zestawów wystąpi nieprawidłowe działanie modułu, tj. stwierdzona zostanie obecność błędów, przy założeniu, że w poprzednich cyklach procesu testowania wykryto i usunięto z modułu $\sum_{i=1}^{k-1} n_i$ błędów. Oczywiście zachodzi:

$$\sum_{m_k=0}^{L_k} A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k) =$$

$$= \sum_{m_k=0}^{L_k} \binom{L_k}{m_k} [e^{-\alpha \sum_{i=1}^{k-1} n_i} (1-r)]^{m_k} [1-e^{-\alpha \sum_{i=1}^{k-1} n_i} (1-r)]^{L_k - m_k} =$$

$$= [1 - e^{-\alpha \sum_{i=1}^{k-1} n_i} (1-r) + e^{-\alpha \sum_{i=1}^{k-1} n_i} (1-r)]^{L_k} = 1$$

oraz

$$\sum_{n_k=0}^{L_k}\sum_{m_k=0}^{L_k} p_{n_k m_k} A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k) = \sum_{m_k=0}^{L_k} A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k)\sum_{n_k=0}^{L_k} p_{n_k m_k} =$$

$$= \sum_{m_k=0}^{L_k} A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k) = 1$$

Wykorzystując zależności (9) i (34) wyrażenie (28), określające wskaźnik niezawodności modułu *r(S,P)* po zakończeniu procesu jego testowania realizowanego według strategii *S*, można przedstawić w następującej postaci:

$$r(S,P) =$$

$$= \sum_{n=0}^{L(S)} [1-(1-r)e^{-\alpha n}] \sum_{n_1+n_2+\ldots+n_K=n} \prod_{k=1}^{K} \sum_{m_k=0}^{L_k} p_{n_k m_k} A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k). \quad (35)$$

Przekształcając zależność (35) otrzymuje się:

$$r(S,P) = \sum_{n=0}^{L(S)} [1-(1-r)e^{-\alpha n}] \sum_{n_1+n_2+\ldots+n_K=n} \prod_{k=1}^{K} \sum_{m_k=0}^{L_k} p_{n_k m_k} A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k) =$$

$$= \sum_{n=0}^{L(S)} \sum_{n_1+n_2+\ldots+n_K=n} \prod_{k=1}^{K} \sum_{m_k=0}^{L_k} p_{n_k m_k} A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k) +$$

$$-(1-r)\sum_{n=0}^{L(S)} e^{-\alpha n} \sum_{n_1+n_2+\ldots+n_K=n} \prod_{k=1}^{K} \sum_{m_k=0}^{L_k} p_{n_k m_k} A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k) =$$

$$= 1-(1-r)\sum_{n_1=0}^{L_1}\sum_{n_2=0}^{L_2}\ldots\sum_{n_K=0}^{L_K} e^{-\alpha\sum_{k=1}^{K} n_k} \prod_{k=1}^{K} \sum_{m_k=0}^{L_k} p_{n_k m_k} A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k),$$

czyli

$$r(S,P) = 1-(1-r)A(S,P), \quad (36)$$

gdzie

$$A(S,P) = \sum_{n_1=0}^{L_1}\sum_{n_2=0}^{L_2}\ldots\sum_{n_K=0}^{L_K} e^{-\alpha\sum_{k=1}^{K} n_k} \prod_{k=1}^{K} \sum_{m_k=0}^{L_k} p_{n_k m_k} A_{m_k}(\sum_{i=1}^{k-1} n_i, L_k). \quad (37)$$

Przyrost wartości wskaźnika niezawodności modułu $\Delta r(S,P) = r(S,P) - r$, uzyskany w wyniku jego testowania realizowanego według strategii *S*, wyraża się następującą zależnością:

$$\Delta r(S,P) = (1-r)(1-A(S,P)), \quad (38)$$

przy czym wielkość *A(S,P)* jest określona zależnością (37).

W przypadku, gdy zbiory zestawów danych wykorzystywane w cyklach procesu testowania modułu są jednoelementowe, tzn. $L_k = 1,\ k = \overline{1,K},$ występujące w zależności (37) prawdopodobieństwa $p_{n_k m_k}$, $k = \overline{1,K}$, określone są następująco:

$$p_{n_k m_k} = \begin{cases} 1 & dla\ n = 0 \ \ i \ \ m = 0\ lub\ n = 1 \ \ i \ \ m = 1, \\ 0 & w\ przeciwnym\ przypadku. \end{cases} \tag{39}$$

W omawianym przypadku strategia testowania ma postać:

$$S' = (L, \underbrace{(1,1,...,1)}_{L\ razy}), \tag{40}$$

a wskaźnik niezawodności modułu, po zakończeniu procesu jego testowania, określony jest następująco:

$$\begin{aligned} & r(S',P) = r(S') = 1-(1-r)\cdot \\ & \sum_{n_1=0}^{1} e^{-\alpha n_1}(1-r)^{n_1} r^{1-n_1} \sum_{n_2=0}^{1} e^{-\alpha n_2}[e^{-\alpha n_1}(1-r)]^{n_2}[1-e^{-\alpha n_1}(1-r)]^{1-n_2} \cdots \\ & \cdots \sum_{n_K=0}^{1} e^{-\alpha n_K}[e^{-\alpha \sum_{m=1}^{K-1} n_m}(1-r)]^{n_K}[1-e^{-\alpha \sum_{m=1}^{K-1} n_m}(1-r)]^{1-n_K}. \end{aligned} \tag{41}$$

Z zależności (41) wynika, że dla strategii S' wartość wskaźnika niezawodności modułu nie zależy od macierzy charakterystycznej P. Zgodnie z przyjętym określeniem strategii testowania (4) strategia S' odpowiada przypadkowi, w którym lokalizacja i usuwanie błędu następuje niezwłocznie po jego wykryciu.

Drugim, charakterystycznym przypadkiem strategii testowania modułu jest strategia obejmująca tylko jeden cykl, tj. strategia postaci

$$S'' = (1, L), \quad L > 0. \tag{42}$$

Wówczas:

$$r(S'',P) = \sum_{n=0}^{L} e^{-\alpha n} \sum_{m=0}^{L} p_{nm} A_m(0,L), \tag{43}$$

czyli, uwzględniając, że:

$$A_m(0,L) = \binom{L}{m}(1-r)^m r^{L-m}$$

otrzymuje się:

$$r(S'',P) = \sum_{n=0}^{L} e^{-\alpha n} \sum_{m=0}^{L} p_{nm} \binom{L}{m}(1-r)^m r^{L-m} . \tag{44}$$

Podstawowa zaleta praktycznego stosowania pierwszej z wymienionych strategii szczególnych, tj. strategii S', wiąże się z wykluczeniem możliwości wykrywania przez różne zestawy danych testowych tego samego błędu (przy założeniu, że raz wykryty błąd jest skutecznie usuwany). Pozwala to zwiększyć skuteczność procesu testowania, mierzoną np. stosunkiem liczby użytych zestawów danych testowych do łącznej liczby wykrytych błędów. Praktyczne stosowanie omawianej strategii może okazać się jednak bardzo kosztowne, z uwagi na wysoki udział kosztów wykonywania prac związanych z lokalizacją i usuwaniem wykrytych błędów w ogólnym koszcie realizacji etapu testowania.

Druga z wymienionych wcześniej strategii szczególnych, tj. strategia S'', z oczywistych względów na ogół nie może stanowić wystarczającej, racjonalnej strategii testowania, np. z uwagi na bardzo często występujący w praktyce efekt tzw. maskowania błędów, polegający na „przesłanianiu" jednych błędów przez inne (co oznacza, że wykrycie pewnych błędów musi być poprzedzone wcześniejszym wykryciem i usunięciem błędów je przesłaniających). Strategia ta stanowi natomiast podstawowa strategię testowania, wykorzystywaną w ramach poszczególnych cykli procesu testowania, przy czym wartość prawdopodobieństwa poprawnego, pojedynczego wykonania się testowanego modułu w poszczególnych cyklach ulega zmianie (w sposób niemalejący - przy założeniu, że wykonanie cyklu testowania modułu nie pogarsza jego niezawodności).

5. Oszacowanie wpływu procesu testowania na wartość wskaźnika niezawodności modułu

Niech $P^*, P^{**} \in \boldsymbol{P}$ będą macierzami charakterystycznymi określonymi następująco:

$$P^{*}=\begin{bmatrix}1 & 0 & 0 & 0 & \dots\\ 0 & 1 & 0 & 0 & \dots\\ 0 & 0 & 1 & 0 & \dots\\ 0 & 0 & 0 & 1 & \dots\\ & & \dots & & \end{bmatrix} \quad P^{**}=\begin{bmatrix}1 & 0 & 0 & 0 & \dots\\ 0 & 1 & 1 & 1 & \dots\\ 0 & 0 & 0 & 0 & \dots\\ 0 & 0 & 0 & 0 & \dots\\ & & \dots & & \end{bmatrix} \tag{45}$$

Można pokazać [19], że powyższe „skrajne” postaci macierzy charakterystycznej testowanego modułu mogą być wykorzystane do skonstruowania dwustronnego oszacowania przyrostu wartości wskaźnika niezawodności modułu dla dowolnej macierzy charakterystycznej $P \in \boldsymbol{P}$ oraz dla dowolnej strategii testowania $S \in \boldsymbol{S}$. Wspomniane oszacowanie jest postaci

$$\Delta r(S,P^{**}) \leq \Delta r(S,P) \leq \Delta r(S,P^{*}) . \tag{46}$$

Warto podkreślić, że zależność (46) może stanowić podstawę do określenia zakresów zmienności wymienionych wielkości dla dowolnej strategii testowania *S*. Oszacowanie to może być szczególnie przydatne w sytuacji, gdy nie jest znana macierz charakterystyczna testowanego modułu *P*.

Wykorzystywane w oszacowaniu (46) macierze charakterystyczne P^{*} i P^{**} odpowiadają specyficznym, tzw. skrajnym schematom testowania. Macierz P^{*} odpowiada bowiem takiemu schematowi, w którym – w ramach każdego cyklu – różne zestawy danych testowych mogą wykrywać jedynie nowe, tj. nie wykryte przez wcześniejsze zestawy danych, błędy. Z kolei, macierz P^{**} odpowiada schematowi, w którym wszystkie zestawy danych, wykorzystywane w ramach określonego cyklu testowania, mogą wykryć co najwyżej jeden i ten sam błąd.

Z uwagi na ww. własności macierzy P^{*}, P^{**} można stwierdzić, że dla określonej strategii testowania modułu $S \in \boldsymbol{S}$ przyrost wartości wskaźnika niezawodności modułu $\Delta r(S,P^{*})$ może być traktowany jako optymistyczny, podczas gdy przyrost $\Delta r(S,P^{**})$ - jako pesymistyczny.

Niech $S', S'' \in \boldsymbol{S}$ oznaczają strategie testowania modułu, określone zależnościami (40) i (42) odpowiednio.

Wówczas dla dowolnej strategii testowania modułu $S \in \boldsymbol{S}$, określonej zależnością (4), zachodzi [17,18]:

$$\Delta r(S',P^{*}) \leq \Delta r(S,P^{*}) \leq \Delta r(S'',P^{*}) \tag{47}$$

oraz

$$\Delta r(S'',P^{**}) \leq \Delta r(S,P^{**}) \leq \Delta r(S',P^{**}) . \tag{48}$$

Zgodnie z określeniem (45) postać macierzy charakterystycznej P^* odpowiada przypadkowi, gdy wśród dowolnej liczby błędów wykrytych w wyniku wykonania określonego cyklu procesu testowania modułu nie występują błędy powtarzające się, tzn. że każdy wykorzystywany zestaw danych testowych może wykryć jedynie nowy, tj. nie wykryty przez wcześniejsze zestawy danych, błąd. Wykorzystując wprowadzone oznaczenia można stwierdzić, że przypadek macierzy charakterystycznej P^* odpowiada sytuacji, gdy dla każdego cyklu procesu testowania modułu, realizowanego według określonej strategii *S,* zachodzi: $M_k(S,P^*) = N_k(S,P^*),\ k = \overline{1,K}$.

Z oszacowania (47) wynika, że w przypadku, gdy macierz charakterystyczna testowanego modułu jest postaci P^* , to spośród wszystkich strategii, testowania spełniających warunek

$$\sum_{k=1}^{K} L_k = L , \tag{49}$$

największy przyrost wartości wskaźnika niezawodności modułu (36) zapewnia strategia *S″*, określona zależnością (40). Otrzymany wynik jest bezpośrednią konsekwencją przyjęcia założeń określających rozpatrywany schemat testowania oraz postaci macierzy P^*, a także natury rozkładu dwumianowego, wykorzystanego do opisu rozkładu prawdopodobieństwa liczby błędów spodziewanych w wyniku wykonania określonego cyklu procesu testowania modułu. Z kolei, jeśli macierz charakterystyczna testowanego modułu jest postaci P^{**} , to spośród wszystkich strategii testowania spełniających warunek (49), największy przyrost wartości wskaźnika niezawodności modułu (36) zapewnia strategia *S″*, określona zależnością (42).

6. Przypadek dwumianowej macierzy charakterystycznej

Macierze charakterystyczne P^*, P^{**} , których wpływ na wartość przyjętego wskaźnika niezawodności testowanego modułu dla strategii *S’,S”* był omówiony w podrozdziale 5, są skrajnymi postaciami macierzy charakterystycznej *P*.
Dla uzupełnienia wyników tej analizy interesujące byłoby zbadanie wpływu „pośredniej” postaci macierzy charakterystycznej na wartości przyrostów wskaźnika niezawodności modułu dla rozpatrywanych strategii *S’,S”*.

Niech macierz charakterystyczna $P \in \boldsymbol{P}$ testowanego modułu spełnia warunek $P = P(\rho)$, przy czym $P(\rho) = [p_{nm}(\rho)]$, gdzie:

$$\mathrm{p}_{\mathrm{nm}}(\rho)=\begin{cases}1 & dla\ \mathrm{n}=\mathrm{m}=0\\ \begin{pmatrix}\mathrm{m\text{-}1}\\ \mathrm{m\text{-}n}\end{pmatrix}\rho^{m-n}(1-\rho)^{n-1} & dla\ \ n=\overline{1,m-1},\ \ m=\overline{1,L_k},\ \ k=\overline{1,K}\\ 0 & dla\ \ n=0\ \ i\ \ \mathrm{m}>0\ \ \mathrm{lub}\ \ n>m\end{cases} \tag{50}$$

przy czym ρ oznacza prawdopodobieństwo zdarzenia polegającego na tym, że kolejny błąd, wykryty w procesie testowania będzie błędem powtarzającym się, tzn. błędem, który już został wykryty przez jeden z wcześniejszych zestawów danych, $0 \leq \rho \leq 1$.

Łatwo sprawdzić, że dla każdej kolumny tak określonej macierzy $P(\rho)$ zachodzi:

$$\sum_{n=0}^{\infty} p_{nm}(\rho)=1, \quad m \in \{0,1,2,...\}. \tag{51}$$

Istotnie, powyższa równość jest prawdziwa, gdyż zgodnie z (50) otrzymuje się:
dla $m=0$:

$$\sum_{n=0}^{\infty} p_{n0}(\rho)=p_{00}(\rho)=1,$$

dla $m>0$:

$$\sum_{n=0}^{\infty} p_{nm}(\rho)=\sum_{n=1}^{m}\begin{pmatrix}m\text{ - }1\\ m\text{ - }n\end{pmatrix}\rho^{m-n}(1-\rho)^{n-1}=\sum_{k=0}^{m-1}\begin{pmatrix}m\text{ - }1\\ k\end{pmatrix}(1-\rho)^{k}\rho^{m-k-1}=1.$$

Z określenia (50) macierzy $P(\rho)$ wynika, że zachodzi:

$$P(0)=P^{*},$$
$$P(1)=P^{**},$$

gdzie macierze $P^{*}, P^{**} \in \boldsymbol{P}$ mają postać określoną zależnością (45).

Niech $S', S'' \in \boldsymbol{S}$ oznaczają strategie testowania modułu, określone zależnościami (40) i (42) odpowiednio, tj.:

$$S'=(L,\ \underbrace{(1,1,...,1)}_{L\ razy}),\quad S''=(1,L),\ \ L>0.$$

Przyjmując, że poszczególne elementy macierzy charakterystycznej $P(\rho)$ określone są zależnością (50), przeanalizowany zostanie obecnie wpływ wartości prawdopodobieństwa ρ na różnicę $A(S'',P(\rho))-A(S',P(\rho))$, a tym samym i na różnicę $\Delta r(S'',P(\rho))-\Delta r(S',P(\rho))$. Wielkość $A(S'',P(\rho))$, określoną zależnością (37), można przedstawić w postaci:

$$A(S'',P(\rho)) = \sum_{n=0}^{L} e^{-\alpha n} \sum_{m=0}^{L} p_{nm}(\rho)\, A_m(0,L)\,, \tag{52}$$

gdzie:

$$A_m(0,L) = \binom{L}{m}(1-r)^m\, r^{L-m}\,. \tag{53}$$

Uwzględniając (50) oraz (53) zależność (52) można przekształcić następująco:

$$\begin{aligned}
A(S'',P(\rho)) &= \sum_{n=0}^{L} e^{-\alpha n} \sum_{m=0}^{L} p_{nm}(\rho) \binom{L}{m}(1-r)^m r^{L-m} = \\
&= \sum_{m=0}^{L} \binom{L}{m}(1-r)^m r^{L-m} \sum_{n=0}^{m} e^{-\alpha n} p_{nm}(\rho) = \\
&= r^L + \sum_{m=1}^{L} \binom{L}{m}(1-r)^m r^{L-m} \sum_{n=1}^{m} e^{-\alpha n} \binom{m-1}{m-n} \rho^{m-n}(1-\rho)^{n-1} = \\
&= r^L + \sum_{m=1}^{L} \binom{L}{m}(1-r)^m r^{L-m} \sum_{k=0}^{m-1} e^{-\alpha} e^{-\alpha k} \binom{m-1}{k} \rho^{m-k-1}(1-\rho)^{k} = \\
&= r^L + e^{-\alpha} \sum_{m=1}^{L} \binom{L}{m}(1-r)^m r^{L-m} [\rho + e^{-\alpha}(1-\rho)]^{m-1},
\end{aligned}$$

czyli

$$A(S'',P(\rho)) = r^L + e^{-\alpha} \sum_{m=1}^{L} \binom{L}{m}(1-r)^m\, r^{L-m}\, [\,\rho + e^{-\alpha}(1-\rho)]^{m-1}. \tag{54}$$

Po dalszych przekształceniach otrzymuje się:

$$\begin{aligned}
&A(S'',P(\rho)) = \\
&= r^L + e^{-\alpha}[\,\rho + e^{-\alpha}(1-\rho)]^{-1} \sum_{m=1}^{L} \binom{L}{m} [\,\rho + e^{-\alpha}(1-\rho)]^{m}(1-r)^m r^{L-m} = \\
&= r^L + e^{-\alpha}[\,\rho + e^{-\alpha}(1-\rho)]^{-1} \{[\,r + (\,\rho + e^{-\alpha}(1-\rho))(1-r)]^L - r^L\} = \\
&= r^L + e^{-\alpha}[\,\rho + e^{-\alpha}(1-\rho)]^{-1} [\,r + (\,\rho + e^{-\alpha}(1-\rho))(1-r)]^L + \\
&\qquad\qquad - r^L e^{-\alpha}[\,\rho + e^{-\alpha}(1-\rho)]^{-1} = \\
&= e^{-\alpha}[\,\rho + e^{-\alpha}(1-\rho)]^{-1} [\,r + (\,\rho + e^{-\alpha}(1-\rho))(1-r)]^L + \\
&\qquad\qquad + r^L \{1 - e^{-\alpha}[\,\rho + e^{-\alpha}(1-\rho)]^{-1}\}.
\end{aligned}$$

Zatem

$$\begin{aligned} &A(S'',P(\rho)) = \\ &= e^{-\alpha}[\rho + e^{-\alpha}(1-\rho)]^{-1}[r + (\rho + e^{-\alpha}(1-\rho))(1-r)]^{L} + \\ &+ r^{L}[1 - e^{-\alpha}[\rho + e^{-\alpha}(1-\rho)]^{-1}]. \end{aligned} \tag{55}$$

Z kolei, zgodnie z (37), wielkość $A(S',P(\rho))$ określona jest następująco:

$$A(S',P(\rho)) = \sum_{n_1=0}^{1}\sum_{n_2=0}^{1}\ldots\sum_{n_L=0}^{1} e^{-\alpha\sum_{l=1}^{L} n_l} \prod_{l=1}^{L}\sum_{m_k=0}^{1} A_{n_l}(\sum_{i=1}^{l-1} n_i, 1),$$

czyli, uwzględniając (34):

$$\begin{aligned} &A(S',P(\rho)) = \\ &= \sum_{n_1=0}^{1} e^{-\alpha n_1}(1-r)^{n_1} r^{1-n_1} \sum_{n_2=0}^{1} e^{-\alpha n_2}[e^{-\alpha n_1}(1-r)]^{n_2}[1-e^{-\alpha n_1}(1-r)]^{1-n_2} \cdots \\ &\cdots \sum_{n_L=0}^{1} e^{-\alpha n_L}[e^{-\alpha\sum_{m=1}^{L-1} n_m}(1-r)]^{n_L}[1-e^{-\alpha\sum_{m=1}^{L-1} n_m}(1-r)]^{1-n_L}. \end{aligned} \tag{56}$$

Niech $S',S'' \in \boldsymbol{S}$ oznaczają strategie testowania modułu postaci (40) i (42) odpowiednio. Niech $P(\rho) = [p_{nm}(\rho)] \in \boldsymbol{P}$, $\rho \in [0,1]$ oznacza macierz charakterystyczną testowanego modułu, określoną zależnością (50). Wówczas istnieje $\rho^* \in (0,1)$, takie że:

$$\forall(\rho < \rho^*)\ (\Delta r(S',P(\rho)) \le \Delta r(S,P(\rho)) \le \Delta r(S'',P(\rho))) \tag{57}$$

oraz

$$\forall(\rho \ge \rho^*)\ (\Delta r(S'',P(\rho)) \le \Delta r(S,P(\rho)) \le \Delta r(S',P(\rho))), \tag{58}$$

przy czym wartość ρ^* jest rozwiązaniem równania:

$$\begin{aligned} &e^{-\alpha}[\rho + e^{-\alpha}(1-\rho)]^{-1}[r + (\rho + e^{-\alpha}(1-\rho))(1-r)]^{L} + \\ &\qquad + r^{L}[1 - e^{-\alpha}[\rho + e^{-\alpha}(1-\rho)]^{-1}] = \\ &= \sum_{n_1=0}^{1} e^{-\alpha n_1}(1-r)^{n_1} r^{1-n_1} \sum_{n_2=0}^{1} e^{-\alpha n_2}[e^{-\alpha n_1}(1-r)]^{n_1}[1-e^{-\alpha n_1}(1-r)]^{1-n_1} \cdots \\ &\cdots \sum_{n_L=0}^{1} e^{-\alpha n_L}[e^{-\alpha\sum_{m=1}^{L-1} n_m}(1-r)]^{n_L}[1-e^{-\alpha\sum_{m=1}^{L-1} n_m}(1-r)]^{1-n_L}. \end{aligned} \tag{59}$$

Zależności (55) i (56) mogą być traktowane jako naturalne uzupełnienie wcześniejszych oszacowań (46)–(48). Dla przyjętej, dwumianowej postaci macierzy charakterystycznej $P(\rho)$, wskazuje one bowiem na istnienie progowej wartości prawdopodobieństwa powtarzalności błędów, poniżej której lepszą strategią testowania jest strategia *S''*, powyżej której natomiast lepszą okazuje się być strategia *S'*, przy czym jedynym kryterium porównywania obu, ww. strategii jest wartość przyrostu wykorzystywanego wskaźnika niezawodności modułu.

Warto podkreślić, że wyniki omawianego oszacowania są całkowicie zgodne z intuicją. W przypadku bowiem, gdy wartość ρ jest bliska zeru, tzn. że prawdopodobieństwo występowania zjawiska powtarzalności błędów jest bardzo małe, intuicyjnie najlepszą strategią testowania powinna być strategia, dająca najlepsze wyniki w sytuacji braku możliwości wykrywania przez różne zestawy danych tego samego błędu, czyli strategia *S''*. Odpowiednio, gdy wartość ρ jest bliska jedności, co oznacza, że prawdopodobieństwo występowania zjawiska powtarzalności błędów jest duże, najlepszą strategią testowania powinna być strategia, która w takiej sytuacji gwarantuje największy przyrost wartości wykorzystywanego wskaźnika niezawodności modułu, czyli strategia *S'*.

Przykład

Niech strategie $S', S'' \in \boldsymbol{S}$, określone zależnościami (40) i (42) odpowiednio, są postaci: $S'=(2, (1,1))$ oraz $S''=(1, 2)$. Stosownie do (55) i (56) otrzymuje się:

$$\begin{aligned}
A(S'',P(\rho)) = {} & e^{-\alpha}[\rho + e^{-\alpha}(1-\rho)]^{-1}[r + (\rho + e^{-\alpha}(1-\rho))(1-r)]^2 + \\
& + r^2[1 - e^{-\alpha}[\rho + e^{-\alpha}(1-\rho)]^{-1}] = \\
& = e^{-\alpha}[\rho + e^{-\alpha}(1-\rho)]^{-1}\{r^2 + 2r[\rho + e^{-\alpha}(1-\rho)](1-r) + \\
& + [\rho + e^{-\alpha}(1-\rho)]^2(1-r)^2\} + r^2\{1 - e^{-\alpha}[\rho + e^{-\alpha}(1-\rho)]^{-1}\} = \\
& = 2r(1-r)e^{-\alpha} + e^{-\alpha}(1-r)^2[\rho + e^{-\alpha}(1-\rho)] + r^2 = \\
& = r^2 + e^{-\alpha}(1-r)\{2r + (1-r)[\rho + e^{-\alpha}(1-\rho)]\}
\end{aligned}$$

oraz

$$A(S',P(\rho)) = r^2 + e^{-\alpha}(1-r^2) - e^{-2\alpha}(1-r)^2(1-e^{-\alpha}).$$

Ponieważ

$$\frac{d[A(S'',P(\rho)) - A(S',P(\rho))]}{d\rho} = e^{-\alpha}(1-r)^2(1-e^{-\alpha}) \geq 0,$$

a rozwiązaniem równania

$$A(S'',P(\rho)) - A(S',P(\rho)) = 0$$

jest wartość $\rho^* = 1 - e^{-\alpha}$, więc dla $\rho < 1 - e^{-\alpha}$ zachodzi:

$$A(S'',P(\rho)) < A(S',P(\rho))$$

czyli

$$\Delta r(S'', P(\rho)) > \Delta r(S', P(\rho)),$$

oraz dla $\rho \geq 1 - e^{-\alpha}$ zachodzi:

$$A(S'', P(\rho)) \geq A(S', P(\rho)),$$

czyli

$$\Delta r(S'', P(\rho)) \leq \Delta r(S', P(\rho)).$$

■

7. Podsumowanie

Skonstruowany w artykule probabilistyczny model wzrostu niezawodności modułu w procesie jego testowania, w odróżnieniu od większości modeli opisywanych w literaturze, uwzględnia możliwość wykrywania przez różne zestawy danych testowych, wykonywanych w ramach określonego cyklu testowania, tego samego błędu. Głównym środkiem umożliwiającym uwzględnienie tego zjawiska jest macierz charakterystyczna testowanego programu *P*.

Podstawowym warunkiem praktycznego wykorzystania proponowanego modelu, jak również uzyskanych na jego podstawie wyników, jest określenie wartości występujących w wykorzystywanych zależnościach wielkości r, α oraz prawdopodobieństw p_{nm}, tworzących macierz charakterystyczną *P*. Z tego powodu wydaje się, że przydatność uzyskanych wyników może być największa w odniesieniu do takich produktów programowych, w odniesieniu do których najłatwiej jest w praktyce określić wartości wymienionych wielkości. W szczególności, takimi produktami są moduły programowe wyodrębniane bądź tworzone zgodnie z kryteriami określonymi np. w pracy [12], uwzględniającymi m.in. takie aspekty budowy modułu jak rozmiar (mierzony np. liczbą linii kodu źródłowego) oraz złożoność funkcjonalna (mierzona np. liczbą i rodzajem realizowanych funkcji).

Należy podkreślić, że racjonalne określenie wymaganych wartości parametrów r, α oraz macierzy *P* wymaga posiadania bogatych i reprezentatywnych danych faktograficznych, charakteryzujących przebieg – realizowanych w przeszłości – prac związanych z procesem testowania podobnych produktów programowych, z uwzględnieniem specyfiki testowanego oprogramowania oraz istniejących uwarunkowań technologiczno-organizacyjnych. Wspomniana baza faktograficzna stanowi ważny element tzw. inżynierskiego warsztatu produkcji oprogramowania. Jej prowadzenie i utrzymywanie jest szczególnie zasadne przy dziedzinowo ustabilizowanej produkcji oprogramowania.

Literatura

[1] Boland P.J., Singh H.: *A Birth-Process Approach to Moranda's Geometric Software-Reliability Model*. IEEE Transactions on Reliability. Vol. 52, 2003.

[2] Chin-Yu H., Lyu M.R., Sy-Yen K.: *A Unified Scheme of Some Nonhomogenous Poisson Process Models for Software Reliability Estimation*. IEEE Transactions on Software Engineering. Vol. 29, 2003.

[3] Cohen D.M., Dalal S.R., Fredman M.L., Patton G.C.: *The AETG System: An Approach to Testing Based on Combinatorial Design*. IEEE Transactions on Software Engineering. Vol.23, No.7, 1997.

[4] Goseva-Popstojanova K., Trivedi K.S.: *Failure Correlation in Software Reliability Models*. IEEE Transactions on Reliability. Vol. 49, 2000.

[5] Jelinski Z., Moranda P.B.: *Software Reliability Research*. Statistical Computer Performance Evaluation. Academic Press, New York, 1972.

[6] Kann S.H.: *Metrics and Models in Software Quality Engineering*. Addison Wesley, 2002.

[7] Kit E.: *Software Testing in Real Word. Improving the Process*. ACM Press, 1995.

[8] Lyu M.R.; Rangarajan S., van Moorsel A.P.A.: *Optimal Allocation of Test Resources for Software Reliability Growth Modeling in Software Development*. IEEE Transactions on Reliability. Vol. 51, 2002.

[9] Malaiya Y.K., Li. M.N., Bieman J.M., Karcich R.: *Software Reliability Growth With Test Coverage*. IEEE Transactions on Reliability. Vol. 51, 2002.

[10] Padberg F.: *Maximum Likelihood Estimates for the Hypergeometric Software Reliability Model*. International Journal of Reliability, Quality & Safety Engineering. Vol. 10, 2003.

[11] Shooman M.L.: *Probabilistic Models for Software Reliability Prediction*. Statistical Computer Performance Evaluation. Academic Press, New York, 1972.

[12] Thayer T.A., Lipov M., Nelson E.C.: *Software Reliability. A Study of Large Project Reality*. North-Holland Publishing Co., New York,1978.

[13] Tian J.: *Better Reliability Assessment and Prediction through Data Clustering*. IEEE Transactions on Software Engineering. Vol. 28, 2002.

[14] Trachtenberg M.: *A General Theory of Software-Reliability Modelling. IEEE Transactions on Reliability*. Vol.39, No.1, 1990.

[15] Vaurio J.K.: *Common-Cause Failure Models, Data, Quantification*. IEEE Transactions on Reliability. 9, Vol. 48, 1999.

[16] Wooff D.A., Goldstein M., Coolen F.P.: *Bayesian Graphical Models for Software Testing*. IEEE Transactions on Software Engineering. Vol. 28, 2002.

[17] Worwa K.: *Estimation of the Program Testing Strategy. Part 1 - The Same Errors Can Not Be Encountered*. Postępy Cybernetyki. Zeszyt 3-4, 1995.

[18] Worwa K.: *Estimation of the Program Testing Strategy. Part 2 - The Same Errors Can Be Encountered*. Postępy Cybernetyki. Zeszyt 3-4, 1995.

[19] Worwa K.: *Ocena przewidywanego poziomu niezawodności programu*. Postępy Cybernetyki. Zeszyt 1-2, 1993.

[20] Yang M.C.K., Chao A.: *Reliability-Estimation & Stopping-Rules for Software Testing Based on Repeated Appearances of Bugs*. IEEE Transactions on Reliability. Vol.44, No.2, 1995.

[21] Zhang X., Pham H.: *Comparisons of Nonhomogeneous Poisson Process Software Reliability Models and its Applications*. International Journal of Systems Science. Vol. 31, 2000.

Recenzent: prof. dr hab. inż. Włodzimierz Kwiatkowski

Praca wpłynęła do redakcji: 26.10.2003r.