Інститут кібернетики імені
Національної академії н
Кам'янець-Подільський націон
імені Івана Огі

МАТЕМАТИЧНЕ ТА К(
МОДЕЛЮВА

Серія: Фізико-матема

Збірник наукових

Випуск 10

ЗМІСТ

УДК 519.21+62

В. Р. Кукурба*, аспірант,
Я. М. Чабанюк**, д-р фіз.-мат. наук, професор

*Національний університет «Львівська політехніка», м. Львів,
**Львівський державний університет безпеки життєдіяльності, м. Львів

ПРОЦЕДУРА СТОХАСТИЧНОЇ ОПТИМІЗАЦІЇ ДЛЯ МОДЕЛІ ТЕСТУВАННЯ З НАПІВМАРКОВСЬКИМИ ПЕРЕКЛЮЧЕННЯМИ

Розглянуто оптимізаційну процедуру для моделі тестування програмного продукту. Стохастичний процес виявлення помилок описаний з допомогою напівмарковського процесу.

Ключові слова: *модель тестування, процедура стохастичної оптимізації, напівмарковський процес.*

Вступ. Дослідження, вдосконалення та аналіз моделей тестування надійності програмного забезпечення (ПЗ) та комп'ютерної техніки в цілому зумовлене створенням нових та вдосконалення існуючих технологій побудови програмних продуктів, розширенням спектру використання автоматизованих систем в сучасному світі, що є передумовою паралельного розвиток усіх складових процесу побудови та впровадження ПЗ в процеси життєдіяльності людини. Підвищення складності та багатокомпонентність сучасних програмних проектів вимагають спеціалізованого підходу під час створення та застосування. В нас час техніка, автоматизовані системи, ПЗ повинні досягати високого рівня надійності, що дозволяє їм ставати ефективним інструментом в світі нових технологій та цілей. Будь яка неполадка може нести за собою серйозні наслідки та втрати, що не припустимо у сучасному світі конкуренції. Саме показник надійності стоїть першим серед не менш важливих показників — якість, живучість, безпека, готовність. Отже збільшення потреб у використанні комп'ютерної техніки у житті людини вимагаю більш строго підходу до визначення рівня надійності продукту застосування. Поняття надійності ПЗ не рідко виділяють окремо [1], оскільки при застосуванні цього поняття до програмних засобів враховують особливості і відмінності цих об'єктів від стандартних (традиційних) технічних систем, для яких в першу чергу розробляється теорія надійності. Першочергова і фундаментальна відмінність програмних проектів від технічних засобів та систем полягає в тому, що програмний продукт не тільки не зношується з часом, що відбувається з технікою, а ще й те, що в результаті процесу використання виявляються та усуваються помилки, не говорячи про можливість модернізації шляхом розширення програми за рахунок нових модулів. Також підвищуються вимоги до надійності та витривалості програм, виникають задачі опти-

мізації процесу тестування, якісного прогн
много продукту. Для розв'язування подібн
вання в даний час використовуються моде
виток програмних технологій та програм
потребу розвитку таких моделей. Постають
делей, вдосконалення існуючих шляхом по
зумовлять підвищення ступеня адекватн
об'єктам, а також тестування таких модел
ливою складовою кожної моделі тестуван
критерій достатності процесу тестування,
проектів приймати обгрунтовані рішення п
розробки. В даний час у переважній більш
ники (критерії) носять скоріше якісний та
що є невиправданим при розробці ПЗ ві
Отже пошук та дослідження критерію дос
є актуальною задачею програмної інженері

**1. Зв'язок між розподілом помилок т
ковським процесом.** Розглядаються моделі
сонівський розподіл кількості помилок на
грамного продукту. Такі моделі розглядаю
Jelinski-Moranda [4], Schick-Wolverton [5],
Okumoto, Schneidewind [7] S-подібна модел
гальнена модель негомогенного пуассонівс
ся, що кількість виявлення помилок у мод
вання надійності ПЗ розподілений за законо
на основі кількості помилок над моделлю на
гає в тому, що визначення часу між окрем
куди меншу точність та більшу похибку за
на перших етапах тестування, на відміну
тестовому проміжку. З іншої точки зору дос
ми може мати свої переваги при оцінці кі
кількості виявлених помилок відчутно змен
ботах [4–7] отримано результати, що описуі
також дають змогу отримати оцінку на гра
результати базуються на отриманій статист
ході подальших досліджень виявлено, що з м
помилок до неоднорідного пуассонівського
відповідає неоднорідному експоненційному

Розглянемо випадок, коли кількості
ціях відповідають закону розподілу Пуасс

Нехай $P_n(t)$ — ймовірність того, щ
залежить від неперервного парамтра t. От
суються розподілом Пуассона:

$$P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}. \quad (1)$$

Розглянемо стохастичний процес, представлений випадковим натуральним числом $X(t)$. Приріст $X(t+s) - X(0)$ на інтервалі від 0 до $t+s$ є сумою приростів $X(s) - X(0)$ і $X(t+s) - X(s)$ на відповідних інтервалах. Ми припускаємо, що прирости $X(s) - X(0)$ і $X(t+s) - X(s)$ незалежні, і що розподіл $X(t+s) - X(s)$ залежить тільки від довжини тестового інтервалу а не від його положення на тестовому проміжку(в даному контексті слід розуміти тільки той тестовий проміжок, кількості помилок на якому відповідають закону Пуассона). На віддалених етапах тестування, нас цікавить, не тільки кількість помилок на проміжку, а й час між помилками, тобто виконання умови $X(t+s) - X(t) = 0$, де s час, що минув після виявлення останньої помилки, тобто ймовірність, що час s менший за час до настання наступної помилки. Це відповідає випадку з розподілу Пуассона, коли $n = 0$:

$$P_1(t) = e^{-\lambda t}. \quad (2)$$

У результаті отримано наступну щільність $f(t) = e^{-\lambda t}$ часу між помилками на пізніх етапах тестування, що відповідає експоненційному закону розподілу, який визначає час перебування у станах марковського процесу. Останній результат дає змогу моделювати подальшу поведінку процесу тестування. З моделі отримуються значення параметра λ, яке можна використати для моделювання ланцюга Маркова, що породжує досліджуваний процес.

У випадку напівмарковського процесу розглядається відповідність кількостей помилок на тестових ітераціях неоднорідному пуассонівському розподілу. При цьому закон розподілу часових проміжків між виявленням помилок може описуватися функцією розподілу відмінною від експоненційної.

У цій статті розглядається функція інтенсивності моделі тестування програмного продукту з індексом величини проекту [1].

2. Аналіз математичної моделі тестування програмного продукту з індексом величини проекту. Математична модель надійності програмного забезпечення створюється для оцінки залежності надійності програмного забезпечення від деяких параметрів, пов'язаних з модулями програми на підмножині наборів вхідних даних, за допомогою, яких цей модуль контролюється [1]. До інших таких параметрів відносяться частота помилок, що дає змогу оцінити якість систем реального часу, що працюють в неперервному режимі, і в той же час паралельно отримувати інформацію про надійність програмних про-

дуктів [3]. Розглянута модель відноситься д
кількості помилок. Припускається, що кільк
моделі оцінювання та прогнозування надій
неоднорідним законом Пуассона. Крім то
величини проекту є параметром моделі та в
спериментальних даних і набуває значенн
завжди більший від нуля. Функція інтенсив
для даної моделі [3] має наступний вигляд

$$\lambda(t) = \alpha\beta^{s+1}t^s \exp(-\beta$$

де α — коефіцієнт, що характеризує зага
ПЗ, β — коефіцієнт, що характеризує за
виявлення помилок, s — індекс величину п

Згідно з [1] для (3) визначена функці
помилок до моменту t

$$\mu(t) = \int_0^t \lambda(\tau)d\tau = \alpha\left[s\Gamma_{\beta t}(s) -\right.$$

де $\Gamma_z(p) = \int_0^z t^{p-1}e^{-t}dt$, $Re(p) > 0$, — непов

Зауваження 1. При $s = 1$ маємо з (3) т
лення

$$\lambda(t) = \alpha\beta^2 t \exp(-\beta t$$

$$\mu(t) = \alpha(1 - (1 + \beta t)\exp(-$$

що відповідає S-подібній моделі [7].

Загальна кількість помилок, що містит
кті визначається з (4) при $t \to \infty$

$$\mu(\infty) = \alpha s\Gamma(s),$$

де $\Gamma(s)$ — гама-функція, що також співпадає

Таким чином, згідно з основним припу
кумулятивної функції (4), для знаходженн
будуємо функцію максимальної правдоподі

$$L(\alpha, \beta, s) = \prod_{i=1}^{n} \frac{\left[\mu(t_i) - \mu(t_{i-1})\right]^{m_i}}{m_i!} \text{ex}$$

де m_i, $\left(\sum_{i=1}^{n} m_i = k\right)$, — кількість виявле
$(t_i; t_{i+1}]$, $i = \overline{0, n}$. Отже, аналітичний вигляд п
узагальнити вираз для загальної кількості по
лежить від величини та складності проекту і

моделі. Рівняння (3) та (4) називають моделлю з індексом величини проекту [3]. Особливістю досліджуваної моделі є третій динамічний параметр, який описує індекс величини програмного проекту, що відсутній у всіх існуючих моделей [4-7]. Під величиною проекту розуміють комплексний показник, який корелює з метриками складності коду програмного продукту [1]. Встановлення залежності між індексом величини проекту та метриками складності коду є предметом подальших досліджень. Для визначення параметрів α, β, s побудованої моделі використовується метод максимальної правдоподібності.

Критерій достатності процесу тестування ПЗ. Важливим прикладним аспектом моделі надійності ПЗ є встановлення кількісного критерію достатності процесу тестування програмного продукту, який дозволяє керівникам програмних проектів більш обґрунтовано приймати рішення про виділення ресурсів на тестування та про завершення цього етапу розробки ПЗ. Так, на даний час, у переважній більшості IT компаній такі показники носять скоріше якісний та неформалізований характер на зразок «задоволення замовника», які жодним чином не можна використовувати, наприклад, при розробці ПЗ відповідального призначення. На противагу до параметрів α та β, залежність індексу величини проекту (який відсутній в усіх інших моделях на основі розподілу Пуассона) виявляє чітку особливість, яку можна покласти в основу критерію достатності процесу тестування. Ця особливість полягає в тому, що при переході до пуассонового розподілу кількості виявлення помилок залежність $s(t)$ стає гладкою, а значення наближається до постійної величини. Таку поведінку залежності $s(t)$ можна зрозуміти, врахувавши, що індекс величини проекту (параметр моделі) є основним параметром, що визначає форму і функцію розподілу, а відповідно і щільність ймовірності випадкової величини, яка у нашому випадку є кількістю виявлення помилки. Отже, на пізніх етапах тестування ПЗ, коли корельовано помилки виявлені та усунені, а час виявлення тих помилок, що залишились відповідає пуассоновому розподілу, якісна характеристика розподілу (параметр $s(t)$) вже практично не змінюється, а змінюються в основному кількісні характеристики (параметри α та β). Таким чином, з використанням критерію достатності процесу тестування можна визначити загальну кількість помилок в програмному продукті за допомогою рівняння (6) і порівнявши її з кількістю вже виявлених та виправлених помилок прийняти обґрунтоване рішення про розподіл ресурсів проекту зі створення програмного продукту.

3. Процедура стохастичної оптимізації в задачі тестування програмного продукту. В даному підрозділі розглядається неперервна процедура стохастичної оптимізації для індексу величини програмного продукту при тестуванні на знаходження помилок в матема-

тичній моделі тестування програмного заб
кості помилок [4; 5]. Отримано умови
асимптотичних значень індексу при неповн
виходу кількості помилок на розподіл Пуас

При безпосередній залежності функці
середовища, що описується напівмарков
розглянуто процедуру стохастичної оптимі

$$\frac{du(t)}{dt} = a(t)\nabla_{b(t)}C(u(t); x(t$$

де

$$\nabla_{b(t)}C(u(t); x(t)) := \frac{C(u(t)+b(t); x(t)) -}{2b}$$

а $x(t), t \geq 0$, рівномірно ергодичний напівм
рному фазовому просторі станів $(X, \mathbf{X})$ з
$\pi(B), B \in \mathbf{X}$ [10].

У цьому випадку експериментатор
асимптотично оптимальний результат наві
гресії, та при всіх значеннях напівмарковсь
10]. Доведення збіжності процедури стох
ймовірністю 1, а саме

$$P\{\lim_{t \to \infty} u(t) = u_0\} = 1$$

проводиться з використанням асимптотичн
ра двокомпонентного марковського проце
$\varepsilon > 0$ — малий параметр серій, розв'язк
збурення та застосування теореми Невел
Теорема 2.8.1, стор. 100).

Отже, важливим кроком застосуванн
оптимізації (7) є встановлення природи сто
функцію регресії $C(u; x)$.

Основною метою роботи є встановлен
ня процедури стохастичної оптимізації (7) д
них значень індексу величини програмного
сивності виявлених помилок, що запропонов
даного підрозділу досліджуються стохаст
оцінок запропонованої моделі опису процес
продукту, а в другій — проводиться аналіз
дури стохастичної оптимізації (7) для індекс

4. Процедура стохастичної оптиміза проекту. Оскільки α, β та s в (3) визн
Пуассона кількості помилок на сусідніх ін

ливість побудови процедури стохастичної оптимізації для них. Лінійність функції інтенсивності $\lambda(t)$ відносно α не дозволяє отримати таку процедуру. Тому $\hat{\alpha}$, отримане з моделі є постійним для процедури стохастичної оптимізації відносно параметрів β та s. З іншої сторони з природи параметру β в (3) слідує, що він змінюється мало при наборі достатньої статистики помилок при тестуванні (перехід кількості виявлених помилок на сусідніх інтервалах $(t_i;t_{i+1}],(t_{i+1};t_{i+2}]$ до розподілу Пуассона), що підтверджують експериментальні данні.

Таким чином, отримуємо єдино можливу процедуру стохастичної оптимізації для параметра s з функцією регресії

$$\lambda(s,t) = \alpha\hat{\beta}^{s+1}t^{s}\exp\left(-\hat{\beta}t\right). \tag{8}$$

Згідно властивостей точкових оцінок $\hat{\alpha}$, $\hat{\beta}$ (при великих n та k) отримуємо їх незміщеність відносно α та β. З іншої сторони α та β змінюються під впливом напівмарковського процесу виявлення помилок.

Отже, отримуємо процедуру стохастичної оптимізації згідно (7) параметра s

$$\frac{ds(t)}{dt} = \frac{a(t)}{2b(t)}\Big[\hat{\alpha}\hat{\beta}^{s(t)+b(t)+1}t^{s(t)+b(t)}e^{-\hat{\beta}t} - \hat{\alpha}\hat{\beta}^{s(t)-b(t)+1}t^{s(t)-b(t)}e^{-\hat{\beta}t}\Big]. \tag{9}$$

Функція інтенсивності виявлення помилок $\lambda(s,t)$ згідно (8) має єдиний максимум.

Теорема. Якщо функції $a(t)>0, b(t)>0$ задовольняють умовам

$$\int_{t_0}^{\infty}a(t)dt = \infty,\ \int_{t_0}^{\infty}a^2(t)dt < \infty,\ \int_{t_0}^{\infty}a(t)b(t)dt < \infty, t_0 > 0, \tag{10}$$

то для процедур стохастичної оптимізації (11) має місце збіжність

$$P\{\lim_{t\to\infty}s(t) = s_0\} = 1, \tag{11}$$

де s_0 таке, що

$$\max\lambda(s,t) = \lambda(s_0,t).$$

Зауваження 2. Відзначимо, що умови (10) з одної сторони дозволяють рухатись випадковій еволюції $u(t)$ до точки $u = u_0$, а з другої сповільнити цей рух таким чином, щоб встигнути досягнути точку $u = u_0$ при $t \to \infty$.

Доведення. Збіжність з ймовірністю 1 (11) слідує з гладкості функції $\lambda(s,t)$, тобто виконання умови Ліпшиця в формі

$$|\nabla_{b(t)}\lambda(s,t) - \lambda_s'(s,t)| \le c$$

умов (10) та твердження теореми збіжності Г

5. Дослідження процедури на реальн
дослідження даних отриманих з реальних т
безпечення. Дані тестування були розбиті н
лок: Trivial, Minor, Major, Critical, Bloker. Т
номірні моменти часу для яких визначалас
ітераційних моментів. Враховуючи різну п
жен тип розглядався окремо і отримано наст

Тип	$\bar{S}$	*S* (ПСО)	В п
Trivial	0,154	0,16	
Minor	0,144	0,201	
Major	0,122	0,296	
Critical	0,139	0,222	
Bloker	0,153	0,166	

У таблиці представлені значення крит
тестування для останнього значення отрим
значається з (6) та граничного отриманого
(9), а також виявлена кількість помилок та
лок, яка визначається з врахування результа
гою процедури та формули (5).

З результатів дослідження видно, що пр
на фінальну стадію, і його необхідно продовж
між отриманими критеріями з моделі та з пр
між виявленою кількістю помилок та очікува
цільність окремого дослідження кожного ти
грамний проект може спеціалізуватися на різ
можуть викликати різну складність проекту ві

Результати дослідження на різних іте
типу помилок Bloker. Помилки цього тип
ітерацій виявлялися в незначних кількостях

Ітерації	$\hat{\alpha}$	$\hat{\beta}$	$\bar{S}$	*S* ПС
800	167.141	0.001725	0.166	0.17
900	165.905	0.001692	0.163	0.17
1000	162.914	0.001645	0.158	0.17
1100	159.694	0.00162	0.156	0.16
1200	156.419	0.001611	0.155	0.16
1300	156.484	0.001589	0.153	0.16

Результати дослідження на різних ітераційних проміжках для типу помилок Minor. Помилки цього типу зявлялися частіше ніж розглянутого вище.

Таблиця 3

Іте-рації	Ви-явле-но	$\bar{\alpha}$	$\bar{\beta}$	$\bar{S}$	*S* ПСО	Помилки модель	По-милки ПСО
800	413	600.035	0.001613	0.155	0.207	558.884	549.837
900	432	611.636	0.001584	0.152	0.206	570.241	560.625
1000	460	639.238	0.001538	0.148	0.205	596.906	586.09
1100	465	631.897	0.00152	0.146	0.203	590.422	579.689
1200	468	621.389	0.001518	0.146	0.202	580.654	570.211
1300	479	624.8	0.0015	0.144	0.201	584.192	573.506

У таблицях показані результати, на кількох проміжних інтервалах, що відображають зміну значень параметрів моделі та граничних значень процедури.

Висновки. Побудовано процедуру стохастичної оптимізації для індексу величини програмного продукту, а саме для параметру s, який є критерієм достатності процесу тестування програмного забезпечення з врахуванням стохастичності процесу тестування. Створена процедура дозволяє оцінити кількість залишкових помилок, що в свою чергу дає можливість передбачити матеріальні витрати на тестування та його доцільність.

Під час досліджень було виявлено, що модель аналізу процесу тестування може бути застосована в реальних умовах при відповідній строгості до документації процесу тестування та на кінцевих етапах розробки програмного коду, бета-тестуваннях. Найефективніше використання розробленої процедури при регресивному тестуванні з використанням автоматизованих тестів.

Складність процесу розробки сучасного програмного забезпечення підвищує складність процесу тестування, що в свою чергу потребує ефективних об'єктів для опису таких процесів. Отже використання напівмарковського процесу в моделях тестування програмного продукту дає переваги в застосуванні таких моделей.

Список використаних джерел:

1. Побудова і дослідження моделі надійності програмного забезпечення з індексом складності проекту / Я. М. Чабанюк, В. С. Яковина, Д. В. Федасюк та ін. // Інженерія програмного забезпечення. Науковий журнал. — К., 2010. — № 1. — С. 24–29.
2. Оптимізація моделі тестування програмного забезпечення з показником величини проекту / Я. М. Чабанюк, В. Р.Кукурба, Л. Б. Гнатів та ін. // Вісник НУ «ЛП» серія «Комп'ютерні науки та
№ 694 — С. 81–89.
3. Оцінювання та прогнозування надійності п
основі моделі з індексом складності проекту
нюк, М. М. Сенів, У. Т. Хімка // Вісник Х
університету. Технічні науки. — 2011. — №
4. Does software reliability growth behavior follo
process / K.-Y. Cai, D.-B. Hu, Ch.-G. Bai, H.
Software Technology. — 2008. — Vol. 50. — P
5. Goel A. L. Software reliability models: assump
bility / A. L. Goel // IEEE Transactions on soft
Vol. SE-11, № 12. — P. 1411–1423.
6. Van Pul M. C. J. Statistical analysis of
M. C. J. Van Pul // CWI Tract. — Amsterdam, I
7. Yamada S. S-shaped reliability growth modellin
S. Yamada, M. Ohba, S. Osaki // IEEE Tr
1983. — Vol. R-32. No. 5. — P. 475–478.
8. Кукурба В. Р. Збіжність одновимірної проце
ції в напівмарковському середовищі / В. Р. К
дна статистика. Актуарна та фінансова мате
цький нац. ун-т. — 2012. — №1. — С. 64–70.
9. Невельсон М. Б. Стохастическая аппроксим
вание / М. Б. Невельсон, Р. З. Хасьминский.
10. Korolyuk V. S. Stochastic Models of Systems
lyuk. — Kluwer : Dordrecht, 1999. — 185 p.

Optimization procedure for testing model
process of errors finding was described using of

Key words: *testing model, stochastic opti
Markov process.*