

О. А. Руденко¹, З. М. Руденко², Г. В. Головка¹, О. Б. Одарущенко³

¹ Полтавський національний технічний університет імені Юрія Кондратюка, Полтава, Україна

² Полтавський коледж нафти і газу Полтавського національного технічного університету імені Юрія Кондратюка, Полтава, Україна

³ Полтавська державна аграрна академія, Полтава, Україна

ЗНАХОДЖЕННЯ ПАРАМЕТРІВ СКОРИГОВАНОЇ ЛІНІЇ ЕКСПОНЕНЦІАЛЬНОЇ АПРОКСИМАЦІЇ ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ ВИЯВЛЕНИХ ДЕФЕКТІВ ПРИ ОЦІНЮВАННІ КІЛЬКОСТІ ВТОРИННИХ ДЕФЕКТІВ ПРОГРАМНИХ ЗАСОБІВ

У статті проведено аналіз місця характеристики надійності програмного забезпечення в структурі моделей якості програмного забезпечення. Визначено, що в ієрархічній структурі більшості моделей якості програмного забезпечення характеристика надійності є першою підхарактеристикою характеристики якості. Виділені п'ять принципів урахування вторинних дефектів програмних засобів. Для урахування вторинних дефектів програмних засобів використовується: теорія динаміки програмних систем, у якій процеси прояву дефектів у програмних засобах розглядаються як результат дії детермінованих потоків дефектів; теорія часових рядів, де виділяються вторинні дефекти із загального потоку дефектів; імітаційне моделювання; модифікація функцій ризику моделей оцінки надійності програмних засобів та функцій, що характеризують параметри цих моделей, внесенням імовірнісних коефіцієнтів; модифікація функцій ризику моделей оцінки надійності програмних засобів шляхом внесення параметра, що визначає число вторинних дефектів, який визначається порівнянням значень полігона частот дефектів з відповідними значеннями функції регресії. Проаналізовано поняття недосконалого відлагодження програмного забезпечення у контексті урахування вторинних дефектів. Обґрунтовано вибір експоненціальної апроксимації полігона частот виявлених дефектів програмних засобів. Наведено приклади моделей оцінки надійності програмних засобів, функції ризику яких містять експоненціальну складову. Розглянуто послідовність знаходження коефіцієнтів функції, одержаної в результаті зміщення лінії експоненціальної апроксимації полігона частот виявлених дефектів програмних засобів. Показано застосування одержаних коефіцієнтів для методики оцінювання числа вторинних дефектів, що ґрунтується на порівнянні даних статистики числа дефектів і даних зміщеної лінії експоненціальної апроксимації полігона частот дефектів. Одержані рівняння скоригованої лінії експоненціальної апроксимації для вибірок малих і великих об'ємів. Одержані формули для обчислення числа вторинних дефектів на часових інтервалах без урахування та із урахуванням поправки Бесселя.

Ключові слова: якість програмного забезпечення, надійність програмного забезпечення, дефект, вторинний дефект, недосконале відлагодження, експоненціальна апроксимація, полігон частот дефектів, модель оцінки надійності програмних засобів, скоригована лінія експоненціальної апроксимації, поправка Бесселя.

Вступ

Розвиток інформаційних технологій та їх широкое впровадження у різні сфери життя людини висуває підвищені вимоги до забезпечення високої якості програмних засобів, що експлуатуються.

Вимоги до програмного забезпечення, що постійно зростають, приводять до появи нових моделей якості програмного забезпечення. В роботі [1] проведено порівняльний аналіз усіх найбільш значущих моделей якості програмного забезпечення за період часу, що складає понад сорок років, та дослідження їх змін. Моделі якості програмного забезпечення мають ієрархічну структуру характеристик та відношень між ними. На першому (вищому) рівні ієрархії знаходиться якість, на другому її характеристики, на третьому підхарактеристики (деякі моделі містять четвертий рівень). Серед розглянутих дев'яти основних моделей якості у шести надійність знаходиться на другому рівні ієрархії [1], а у моделі Дромера [2] є підхарактеристикою всіх характеристик другого рівня ієрархії. Крім того, нова модель якості, що відображає особливості забезпечення і оцінювання якості сучасних програмних засобів, яка описана у міжнародному спеціалізованому стандарті

ISO 25010 [3], містить характеристику надійності на другому рівні ієрархії. Тому надійність є одним з ключових факторів, що забезпечують якість програмного забезпечення. Водночас важливе значення має оцінка надійності програмного забезпечення, що враховує фактори усіх етапів життєвого циклу програмних продуктів. Одним з таких факторів є фактор вторинних дефектів – дефектів, що вносяться у процесі усунення виявлених (первинних) дефектів програмних засобів. При знаходженні показників надійності неврахування фактора вторинних дефектів може привести до значних відхилень одержаних результатів від реальних, або навіть повної неадекватності цих результатів.

Відомі п'ять принципів врахування вторинних дефектів:

- теорія динаміки програмних систем [4-6];
- виділення вторинних дефектів із загального потоку дефектів з використанням теорії часових рядів [7];
- використання імітаційного моделювання [8];
- модифікація функції ризику моделі оцінки надійності програмного засобу шляхом внесення в неї імовірнісних коефіцієнтів, що характеризують вторинні дефекти [9];

– оцінка числа вторинних дефектів на часових інтервалах шляхом порівняння значень тренда дефектів з відповідними значеннями функції регресії для подальшого використання одержаних результатів у модифікованих (шляхом внесення у функцію ризику моделі параметра, що характеризує число вторинних дефектів) моделях оцінки надійності програмних засобів [10].

У ряді робіт [11-13] розроблені моделі, що характеризують параметри недосконалого відлагодження програмного забезпечення. У статті [11] запропоновано імітаційний підхід до моделювання процесу виявлення дефектів і процесу їх корекції. У роботі [12] розглядаються процеси виявлення та усунення дефектів з урахуванням функції тестування зусиль і недосконалого відлагодження. У статті [13] запропонована модель надійності програмного забезпечення, що пов'язує явище недосконалого

відлагодження і навчання загальним параметром між двома функціями. Фактор вторинних дефектів іноді враховується в комплексі з іншими факторами, що визначають недосконале відлагодження, але не виділяється окремо.

Мета статті – визначення коефіцієнтів скоригованої лінії експоненціальної апроксимації статистичних даних числа виявлених дефектів програмних засобів для оцінки числа вторинних дефектів на часових інтервалах.

Основна частина

Вибір експоненціальної апроксимації зумовлений тим, що значна кількість моделей оцінки надійності програмних засобів включає параметри, що описуються за допомогою рівностей, які містять експоненту. Приклади таких моделей наведені у табл. 1.

Таблиця 1 – Моделі, що містять експоненціальну складову

Модель	Формула	Параметри моделі
Муси-Гамільтона [15]	$m = E_0 \left(1 - e^{-C/E_0 T_0}\right)$	m – число відмов (виявлених дефектів); T_0 – напрацювання між відмовами перед початком налагодження або експлуатації, E_0 – початкове число дефектів, C – коефіцієнт пропорційності
Шнайдевінда [16]	$\lambda(t) = B e^{-K t}$	$\lambda(t)$ – інтенсивність відмов; i – порядковий номер інтервалу; $B > 0$, $K > 0$ – константи моделі
Базова S-подібна модель [17]	$\lambda(t) = B K^2 t e^{-K t}$	$\lambda(t)$ – інтенсивність відмов; B , $K > 0$ – параметри моделі
Охба [18]	$\lambda(t) = a b^2 t e^{-b t}$	$\lambda(t)$ – інтенсивність відмов; a – початкове число дефектів у програмному засобі; b – інтенсивність усунення дефектів

Крім того, багато параметрів, що характеризують надійність програмних засобів, обчислюються за допомогою формул, які містять експоненціальну складову.

Надійність систем, що не відновлюються, визначається величинами:

– ймовірність безвідмовної роботи

$$P(t) = e^{-\int_0^t \lambda(t) dt}, \quad (1)$$

де $\lambda(t)$ – інтенсивність відмов – основний показник надійності складних систем;

– середній час безвідмовної роботи

$$T = \int_0^{\infty} P(t) dt \quad (2)$$

та рядом інших величин, котрі безпосередньо пов'язаних із зазначеними [14].

Методика оцінювання числа вторинних дефектів програмних засобів

Вихідними даними для оцінювання числа вторинних дефектів програмних засобів є статистичні дані числа виявлених дефектів.

Створюється дискретний варіаційний ряд частот виявлених дефектів, варіантами якого є кінці

рівновіддалених часових інтервалів, на які розбивається весь час процесу виявлення дефектів.

Оцінювання числа вторинних дефектів програмних засобів здійснюється у наступній послідовності.

1 Будується полігон частот дефектів, що є лананою, яка послідовно сполучає точки з координатами $(t_i; n_i)$, де t_i – час, що відповідає кінцю i -го інтервалу часу; n_i – число дефектів, виявлених за i -й інтервал часу.

2 Проводиться експоненціальна апроксимація одержаного полігона частот і визначається рівняння лінії апроксимації.

3 Підстановкою в одержане рівняння експоненціальної апроксимації t_1 і t_k визначаються точки

$$(t_1; n_1) \text{ і } (t_k; n_k),$$

де t_1 – час, що відповідає кінцю першого інтервалу часу; t_k – час, що відповідає кінцю останнього інтервалу часу; n_1 , n_k – відповідні значення апроксимуючої функції.

4 Визначається точка з координатами

$$(t_1; n_1 - \sigma),$$

де σ – середнє квадратичне відхилення числа виявлених дефектів.

5 Знаходиться рівняння скоригованої лінії експоненціальної апроксимації. Для цього у рівняння експоненти

$$n = \alpha \exp(-\beta t) \quad (3)$$

замість n і t підставляються координати таких точок:

$$(t_1; n_1 - \sigma) \text{ і } (t_k; n_k).$$

У результаті одержується система рівнянь

$$\begin{cases} n_1 - \sigma = \alpha \exp(-\beta t_1) \\ n_k = \alpha \exp(-\beta t_k) \end{cases} \quad (4)$$

Поділивши перше рівняння системи (4) на друге одержуємо

$$\frac{n_1 - \sigma}{n_k} = \exp(\beta(t_k - t_1)) \quad (5)$$

Прологарифмувавши (5) знаходимо β

$$\beta = \frac{1}{t_k - t_1} \ln \frac{n_1 - \sigma}{n_k} \quad (6)$$

Переписавши друге рівняння системи (4) у вигляді

$$\alpha = n_k \exp(\beta t_k) \quad (7)$$

підставимо (6) замість β .

Одержуємо

$$\alpha = n_k \exp \left(t_k \ln \left(\frac{n_1 - \sigma}{n_k} \right)^{\frac{1}{t_k - t_1}} \right). \quad (8)$$

Після спрощень (8) маємо

$$\alpha = n_k \left(\frac{n_1 - \sigma}{n_k} \right)^{\frac{t_k}{t_k - t_1}}. \quad (9)$$

Знайдемо $\exp(-\beta t)$ підставивши вираз (6) замість β .

$$\exp(-\beta t) = \exp \left(- \frac{1}{t_k - t_1} \ln \frac{n_1 - \sigma}{n_k} t \right), \quad (10)$$

або після спрощень

$$\exp(-\beta t) = \left(\frac{n_1 - \sigma}{n_k} \right)^{\frac{t}{t_k - t_1}}. \quad (11)$$

Підставивши в рівняння експоненти (3) вирази (9) і (11) одержуємо рівняння скоригованої лінії експоненціальної апроксимації

$$n = n_k \left(\frac{n_1 - \sigma}{n_k} \right)^{\frac{t_k - t}{t_k - t_1}}. \quad (12)$$

Число вторинних дефектів на кожному часовому інтервалі визначається за формулою

$$n^{BH} = n_i - n_i^*, \quad i = \overline{1, k}, \quad (13)$$

де n_i – число дефектів, виявлених за i -й інтервал часу, n_i^* – значення скоригованої лінії експоненціальної апроксимації у моменти часу t_i .

Або, враховуючи (12)

$$n^{BH} = n_i - n_k \left(\frac{n_1 - \sigma}{n_k} \right)^{\frac{t_k - t_i}{t_k - t_1}}, \quad (14)$$

де t_i ($i = \overline{1, k}$) – час, що відповідає кінцю i -го інтервалу часу.

Результат, обчислений за формулою (14), округлюється до цілих, n^{BH} вважається рівним нулю, якщо отримується від'ємне значення.

Вибір величини середнього квадратичного відхилення як міри відхилення лінії експоненціальної апроксимації від скоригованої лінії експоненціальної апроксимації у момент часу t_1 на величину σ (рис. 1) обумовлюється тим, що середнє квадратичне відхилення характеризує міру розсіяння значень випадкової величини відносно середнього.

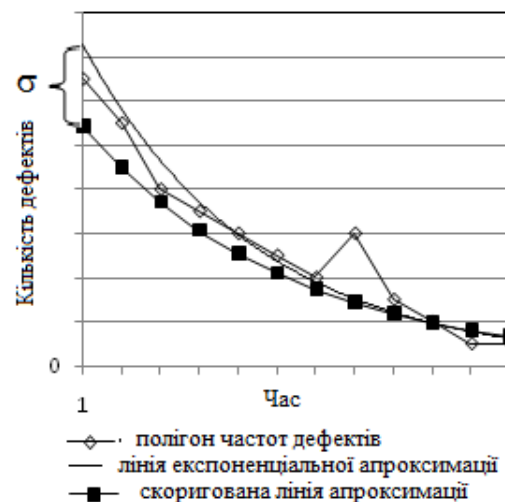


Рис. 1. Полігон частот дефектів, лінія експоненціальної апроксимації та скоригована лінія експоненціальної апроксимації

Зменшення різниці між відповідними значеннями апроксимуючої лінії і скоригованої лінії апроксимації із збільшенням t_i пов'язане з тим, що на початкових етапах тестування програмних засобів більша ймовірність внесення вторинних дефектів, оскільки більше число виявлених дефектів і, відповідно, проводиться більше дій, спрямованих на їх

усунення. У момент часу t_k ці значення співпадають. Це пояснюється припущенням, що на останньому етапі тестування всі виявлені дефекти усунути.

Приклад застосування методики оцінювання числа вторинних дефектів на основі порівняння значень функції експоненціальної апроксимації та функції скоригованої лінії експоненціальної апроксимації наведений у роботі [19] без урахування одержаних формул.

При невеликих вибірках числа дефектів для розрахунку середнього квадратичного відхилення вводять поправку Бесселя. Відповідно, формула (12) набуде вигляду

$$n = n_k \left(\left(n_l - \sqrt{\frac{N}{N-1}} \sigma \right) / n_k \right)^{\frac{t_k - t_l}{t_k - t_l}}, \quad (15)$$

де N – об'єм вибірки.

Формула (14) з урахуванням поправки Бесселя матиме вигляд

$$n^{BH} = n_l - n_k \left(\left(n_l - \sqrt{\frac{N}{N-1}} \sigma \right) / n_k \right)^{\frac{t_k - t_l}{t_k - t_l}}. \quad (16)$$

Для вибірок об'єму $N > 30$ зміщене середнє квадратичне відхилення мало відрізняється від незміщеного, тому в поправці Бесселя немає необхідності [20].

Висновки

Одержані формули для оцінювання числа вторинних дефектів програмних засобів на основі порівняння значень експоненціальної апроксимуючої функції полігона дефектів та зміщеної лінії експоненціальної апроксимації без урахування та з урахуванням поправки Бесселя.

Результати узгоджуються з раніше отриманими результатами за допомогою порівняння значень полігона частот дефектів з відповідними значеннями лінії регресії, але використання одержаного аналітичного виразу значно спрощує обчислювальний процес.

Подальшими напрямками досліджень, спрямованих на оцінку числа вторинних дефектів за статистичними даними є аналіз припущень та аналітичних функцій ризику моделей, що не містять експоненціальної складової та використання інших видів апроксимуючих функцій, а також урахування величини достовірності апроксимації при розрахунках.

СПИСОК ЛІТЕРАТУРИ

1. Гордеев А. А. Эволюция моделей качества программного обеспечения: методика и результаты анализа в контексте стандарта ISO 25010 / А. А. Гордеев, В. С. Харченко // Системы обработки информации. – 2013. – №6(113), С. 15-34.
2. Dromey G. R. A model for software product quality // IEEE Trans. on software Eng. – 1995. Vol.21, no. 2, pp.146-162.
3. International standard ISO/IEC FDIS 25010. System and software quality models. – 2010, 34 p.
4. Маевський Д.А. Структурна динаміка програмних систем і прогнозування їх надійності при наявності вторинних дефектів / Д.А. Маевський // Радіоелектронні і комп'ютерні системи. – 2010. – № 3 (44). – С. 103-109.
4. CASE-оценка критических программных систем. Т. 2. Надежность [Монография] / Одарушенко О.Н., Харченко В.С., Маевский Д.А. и др. – Под ред. Харченко В.С. – Х.: Нац. аэрокосм. ун-т им. Н.Е. Жуковского "ХАИ", 2012. – 292 с.
5. Maevsky D. A. Software reliability. Non-probabilistic approach / D. A. Maevsky, H. D. Maevskaya, A. A. Leonov // RT&A # 03 (26) – 2012, P. 8-20.
6. Маевский Д. А. Использование теории временных рядов для выделения вторичных ошибок на этапе тестирования программного обеспечения / Д.А. Маевский, О.П. Жеков // Радіоелектронні і комп'ютерні системи. – 2011. – № 2 (16). – С. 82-85.
7. Мищенко В.О. Термодинамический подход к моделированию процесса роста надежности программных средств с учетом «вторичных дефектов» / В.О.Мищенко // Вісник Харківського національного університету імені В.Н. Каразіна, серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління». – 2015. – Вип. 28. – С. 91-106.
8. Mahapatra G. S. Modified Jelinski-Moranda Software Reliability Model with Imperfect Debugging Phenomenon / G. S. Mahapatra, P. Roy // International Journal of Computer Applications – 2012. – № 18 – P. 38-46.
9. Одарушенко О.Н. Метод оценивания надежности программных средств с учетом вторичных дефектов / О.Н. Одарушенко, А.А. Руденко, В.С. Харченко // Радіоелектронні і комп'ютерні системи. – 2012. – № 7 (59). – С. 294-300.
10. Peng R., Liu J. Simulated Software Testing Process Considering Debuggers with Different Detection and Correction Capabilities. International Journal of Performability Engineering, Vol. 13, no. 3, 2017, 334-336
11. Peng R, Li YF, Zhang WJ, Hu QP. Testing effort dependent software reliability model for imperfect debugging process considering both detection and correction. Reliability Engineering & System Safety. 2014; 126: 37-43.
12. Pham H. An Imperfect-debugging Fault-detection Dependent-parameter Software. International Journal of Automation and Computing. 2007; 04(4): 325-328.
13. Половко А.М. Основы теории надежности. / А.М. Половко, С.В. Гуров – СГ БХВ-Петербург, 2006. – 704 с.
14. Hamilton P.A., Musa J.D. Measuring reliability of Computation Center Software // Proc.3-thInternat. Conf. on Software. Eng. May 10-12 1978. – P.29-36.
15. Schneidewind N. F. Software Reliability Model with Optimal Selection of Failure Data // IEEE Transactions on Software Engineering. - 1993. - Vol. 19, No. 11. Nov. - P. 1095-1104.
16. Yamada S., Ohba M., Osaki S. S-shaped software reliability grows modeling for software error detection // IEEE Trans. Reliability. - 1983. - R-32. - № 5. - P. 475-518.

17. Ohba M. Software Reliability Models // IBM J. Res. Develop. - 1984. - 28. - № 4. - P. 428-443.
18. Rudenko O., Odarushchenko E., Rudenko Z., Rudenko M., "The Secondary Software Faults Number Evaluation Based on Correction of the Experimental Data Exponential Line Approximation", Conference Proceedings of 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies DESSERT'2018, Kyiv, 2018, pp. 401-405.
19. Мармоза А.Т. Практикум по математической статистике: Учеб. пособие / А.Т. Мармоза – К.: Выща шк., 1990. – 191 с.

Рецензент: д-р техн. наук, проф. В. А. Краснобаев,
Харківський національний університет імені В. Н. Каразіна, Харків

Received (Надійшла) 03.10.2018

Accepted for publication (Прийнята до друку) 28.11.2018

Нахождение параметров скорректированной линии экспоненциальной аппроксимации экспериментальных данных выявленных дефектов при оценке числа вторичных дефектов программных средств

А. А. Руденко, З. Н. Руденко, Г. В. Головки, Е. Б. Одарушенко

В статье проведен анализ места характеристики надежности программного обеспечения в структуре моделей качества программного обеспечения. Определено, что в иерархической структуре большинства моделей качества программного обеспечения характеристика надежности является первой подхарактеристикой характеристики качества. Выделены пять принципов учета вторичных дефектов программных средств. Для учета вторичных дефектов программных средств используется: теория динамики программных систем, в которой процессы проявления дефектов в программных средствах рассматриваются как результат действия детерминированных потоков дефектов; теория временных рядов, где выделяются вторичные дефекты из общего потока дефектов; имитационное моделирование; модификация функций риска моделей оценки надежности программных средств и функций, характеризующих параметры этих моделей, внесением вероятностных коэффициентов; модификация функций риска моделей оценки надежности программных средств путем внесения параметра, определяющего число вторичных дефектов, который определяется сравнением значений полигона частот дефектов с соответствующими значениями функции регрессии. Проанализировано понятие несовершенной отладки программного обеспечения в контексте учета вторичных дефектов. Обоснован выбор экспоненциальной аппроксимации полигона частот выявленных дефектов программных средств. Приведены примеры моделей оценки надежности программных средств, функции риска которых содержат экспоненциальную составляющую. Рассмотрена последовательность нахождения коэффициентов функции, полученной в результате смещения линии экспоненциальной аппроксимации полигона частот выявленных дефектов программных средств. Показано применение полученных коэффициентов для методики оценки числа вторичных дефектов, которая строится на сравнении данных статистики числа дефектов и данных смещенной линии экспоненциальной аппроксимации полигона частот дефектов. Получены уравнения скорректированной линии экспоненциальной аппроксимации для выборок малых и больших объемов. Получены формулы для вычисления числа вторичных дефектов на временных интервалах без учета и с учетом поправки Бесселя.

Ключевые слова: качество программного обеспечения, надежность программного обеспечения, дефект, вторичный дефект, несовершенная отладка, экспоненциальная аппроксимация, полигон частот дефектов, модель оценки надежности программных средств, скорректированная линия экспоненциальной аппроксимации, поправка Бесселя.

Finding the parameters of the corrected line of exponential approximation of the experimental data of the detected faults during estimating the number of secondary faults of the software

O. Rudenko, Z. Rudenko, G. Golovko, O. Odarushchenko

The article analyzes the location of the characteristic reliability of the software in the structure of software quality models. It has been determined that in the hierarchical structure of most software quality models, the reliability characteristic is the first characterization characteristics of quality. There are five principles for recording secondary faults of software. To account for secondary faults of software means: the theory of dynamics of software systems, in which the processes of manifestation of faults in software are considered as the result of the action of deterministic faults streams; the theory of time series, which distinguish secondary faults from the general flow of faults; simulation modeling; modification of the risk functions of models for assessing the reliability of software and functions that characterize the parameters of these models, the introduction of probability coefficients; Modification of the risk functions of software reliability evaluation models by introducing a parameter that determines the number of secondary faults, which is determined by comparing the values of the fault frequency polygon with the corresponding values of the regression function. The concept of imperfect software debugging in the context of accounting for secondary faults is analyzed. The choice of the exponential approximation of the frequency range of detected software faults is grounded. Examples of models for assessing the reliability of software tools whose risk features contain an exponential component are given. The sequence of finding the coefficients of the function obtained as a result of displacement of the exponential approximation line of the frequency range of detected faults of software is considered. The application of the obtained coefficients for the method of estimating the number of secondary faults is shown. This method is based on a comparison of the statistics of the number of faults and the displaced line data of the exponential approximation of the fault frequency polygon. The equations of the corrected exponential approximation line for the samples of small and large volumes are obtained. Formulas have been obtained for calculating the number of secondary faults on time intervals without account and taking into account the Bessel correction.

Keywords: software quality, software reliability, fault, secondary fault, imperfect debugging, exponential approximation, defect frequency polygon, software reliability evaluation model, corrected exponential approximation line, Bessel correction.