

УЧЕТ ФАКТОРА ВТОРИЧНЫХ ДЕФЕКТОВ ПРИ ОЦЕНКЕ НАДЕЖНОСТИ ПРОГРАММНЫХ СРЕДСТВ

В.С. ХАРЧЕНКО¹
А.А. РУДЕНКО²
О.Н. ОДАРУЩЕНКО³
Е.Б. ОДАРУЩЕНКО²

1) Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», г. Харьков, Украина

2) Полтавский национальный технический университет имени Юрия Кондратюка, Украина

3) Научно-производственное предприятие «Радий», г. Кировоград, Украина

*e-mail:
v_s_kharchenko@ukr.net
olekrudenko@yandex.ru
skifs2005@mail.ru
skifs2007@mail.ru*

В статье рассмотрена проблема учета вторичных программных дефектов при оценке надежности программных средств. Проанализированы существующие модели оценки надежности программных средств с точки зрения возможности их использования для оценки надежности программных средств с учетом вторичных дефектов. Предложен метод оценивания надежности программных средств с учетом вторичных дефектов.

Ключевые слова: программное средство, надежность, вторичный дефект программных средств, интенсивность проявления дефектов, модели оценки надежности программных средств.

Введение

Требования к качеству и надежности программных средств (ПС) определяются международными стандартами ISO/IEC 12207, ISO/IEC 9126, ISO/IEC 25010 и др., что находит отображение практически на всех этапах жизненного цикла ПС.

Надежность ПС определяется как набор составляющих, относящихся к «способности программного обеспечения сохранять свой уровень качества функционирования при установленных условиях за установленный период времени»¹.

Надежность ПС является наиболее важной подхарактеристикой качества, наряду с функциональностью, практичностью, эффективностью, сопровождаемостью, мобильностью. Исходя из этого, важной задачей является не только обеспечение высокой надежности ПС, но и ее правильная оценка, что обеспечит минимизацию затрат или рисков в ходе применения их по назначению.

Существует множество моделей оценки надежности ПС (МНПС), группируемых в соответствии с рассматриваемыми в них параметрами по различным классификационным признакам. Существует также множество классификаций МНПС: Хетча, Гоэла, Полонникова-Никандрова, Фатуева, Благодатских и др.². Предложен вариант матричной классификации моделей роста надежности ПС³.

Во всех моделях оценки надежности ПС прямо или косвенно присутствует величина, характеризующая надежность, – дефект (fault). Дефект является основной характеристикой для нахождения показателей при количественной оценке надежности: средней

¹ ISO/IEC 9126-1: 2001. Software engineering – Product quality – Part 1: Quality model.

² Goel A.L. Software Reliability Models: Assumptions, Limitations, and Applicability / Goel A.L. // IEEE Trans. Software Eng., vol. 11, no. 12, 1985. – P. 1411-1423; Haag S. Quality Function Deployment. Usage in Software Development / Haag S., Raja H.K., Sekade L.L. // Comm. of ACM. – Oct., 2002. – P.35-40; Фатуев В.П. Надежность автоматизированных информационных систем: Учебное пособие // В.П. Фатуев, В.И. Высоцкий, В.И. Бушинский – Т.: ТГУ, 1998. – 104 с.; Стандартизация разработки программных средств: Учеб. пособие / В.А. Благодатских, В.А. Волнин, К.Ф. Посакалов; Под ред. О.С. Разумова. – М.: Финансы и статистика, 2005. – 288 с.; Полонников Р.И. Методы оценки показателей надежности программного обеспечения / Р.И. Полонников, А.В. Никандров. – СПб.: Политехника – 1992. – 78 с.

³ Kharchenko V. The Method of Software Reliability Growth Model Choice Using Assumption Matrix / Kharchenko V., Tarasyuk O., Sklyar V., Dubnitsky // Proceedings of the 26th Conference Computer Software and Applications, Oxford, England, 26-29 August, 2002. – P.541-547.



наработки на отказ, вероятности безотказной работы, числа неустранимых программных дефектов, вероятности отсутствия дефектов и др.

Большое количество моделей оценки надежности ПС вызвано разнообразием программных проектов и применяемых технологий, что не позволяет в полной мере отображать реалии проектирования, верификации и сопровождения конкретного ПС. Одной из наиболее важных задач является нахождение рационального сочетания между системой допущений, принятой для соответствующей модели, и уровнем сложности ее математического описания. Существует подход к выбору МНПС, основанный на анализе матриц допущений¹.

В ходе проведенных исследований выявлены основные недостатки формирования допущений для известных моделей:

– предполагается, что обнаруженные (первичные) дефекты устраняются, при этом новые (вторичные) не вносятся, или о возможности внесения дефектов во время устранения обнаруженных речь не идет вообще;

– предполагается экспоненциальный (или другой) закон распределения длительностей временных интервалов обнаружения дефектов и не учитывается время, которое затрачивается на их устранение и др.

Актуальность проблемы вторичных дефектов возрастает с увеличением требований к сложности программного обеспечения. Так, если 30 лет назад доля вторичных дефектов оценивалась в среднем от 2 до 9%², то в настоящее время она составляет более 25%³, что подтверждается мнением экспертов-разработчиков программного обеспечения⁴.

Существует два подхода к оценке надежности ПС с учетом вторичных дефектов.

Первый состоит в модификации существующих МНПС путем введения величины, характеризующей вторичные дефекты⁵. Второй подход, предложенный Д.А. Маевским, строится на основе теории динамики программных систем, где показана и использована аналогия между термодинамической теорией переноса и структурными процессами в программных системах, что позволяет проводить моделирование их надежности⁶.

Цель статьи – рассмотреть аспекты повышения точности оценки надежности ПС на основе учета фактора вторичных дефектов.

1. Классификация вторичных дефектов программных средств.

Сценарии внесения и устранения дефектов

Исследованы работа и сопровождение учетной информационной системы «АгроКомплекс»⁷, которая служит для автоматизации всех видов учета на предприятиях сельского хозяйства. На основе экспериментальных данных сделан вывод, что всплески проявления дефектов соответствуют периодам времени, когда происходило значительное обновление структурных параметров системы⁸. Одной из причин данного отклонения можно считать внесенные дефекты. Кроме того, следует учесть возможность появления дефектов при взаимодействии первичной и обновленной частей ПС.

Таким образом, вторичные дефекты классифицируются следующим образом:

– дефекты, вносимые в процессе устранения первичных дефектов;

¹ Kharchenko V. Указ. соч.

² Тейер Т. Надежность программного обеспечения: Пер. с англ. / Т. Тейер, М. Липов, Э. Нельсон. – М.: Мир, 1981. – 323 с.

³ Совершенный код. Мастер-класс. Пер. с англ. / С. Макконнелл – М.: Издательско-торговый дом «Русская Редакция» ; СПб.: Питер, 2005. – 896 с.

⁴ Маевський Д.А. Структурна динаміка програмних систем та прогнозування їх надійності при наявності вторинних дефектів. // Д.А. Маевський // Радіоелектронні і комп'ютерні системи. – 2010. – №3 – С. 103-109.

⁵ Одарущенко О.Н. CASE-оценка критических программных систем. В 3-х томах. Том 2. Надежность / О.Н. Одарущенко, В.С. Харченко, Д.А. Маевский, Ю.Л. Поночовный, А.А. Руденко, Е.Б. Одарущенко, С.А. Засуха, В.О. Жадан, С.В. Живило / Под ред. Харченко В.С. – Х.: Нац. аэрокосмический ун-т «Харьк. авиац. ин-т», 2012. – 292 с.; Одарущенко О.Н. Учет вторичных дефектов в моделях надежности программных средств / О.Н. Одарущенко, А.А. Руденко, В.С. Харченко // Математичні машини і системи. – 2010. – № 1. – С. 205-217.; Руденко А.А. Модели оценки надежности программных средств с учетом недетерминированного числа вторичных дефектов / А.А. Руденко, О.Н. Одарущенко, В.С. Харченко // Радіоелектронні і комп'ютерні системи. – 2010. – № 6 (47). – С. 197-203.

⁶ Маевський Д.А. Указ. соч.

⁷ Маевський Д.А. Інформаційна система «АгроКомплекс» для бухгалтерського та оперативного обліку у сільському господарстві / Д.А. Маевський, Т.Я. Тінтулова, В.М. Антошук. // Зб. Аграрний вісник причорномор'я. Технічні науки. 2009, – Вип. 48, С. 151-156.

⁸ Маевський Д.А. Структурна динаміка програмних систем...

- дефекты, вносимые при обновлении (в той части ПС, которая обновляется – «условно вторичные дефекты»);
- дефекты взаимодействия первичной и обновленной частей ПС.

На основе приведенной классификации путем перебора возможных вариантов разработаны сценарии внесения и устранения дефектов, характеризующие процессы обнаружения, устранения, внесения дефектов, а также, неустранения обнаруженных дефектов:

$$M_i = \begin{cases} N_{\Sigma} - N_i, \\ N_{\Sigma} - N_i + K_i, \\ N_{\Sigma} - N_i + \Delta N_i, \\ N_{\Sigma} - N_i + K_i + \Delta N_i, \\ N_{\Sigma} - N_i + K_i^*, \\ N_{\Sigma} - N_i + K_i^* + \Delta N_i, \\ N_{\Sigma} - N_i + K_i^* + K_i^B, \\ N_{\Sigma} - N_i + K_i^* + K_i^B + \Delta N_i, \end{cases} \quad (1)$$

где M_i – число дефектов после i -той операции по устранению выявленных дефектов;

N_{Σ} – число дефектов в исходной программе;

N_i – число обнаруженных дефектов;

K_i – число дефектов, внесенных в процессе устранения обнаруженных дефектов;

ΔN_i – число неустраненных дефектов (из числа выявленных);

K_i^* – число дефектов, внесенных в процессе обновления программы;

K_i^B – число дефектов, связанных с взаимодействием первичной и обновленной частей ПО¹.

Выбран сценарий для дальнейшего исследования, базирующийся на соотношении $M_i = N_{\Sigma} - N_i + K_i$. Выбор сделан исходя из анализа допущений МНПС, а также, возможности исследования в системе допущений, принятых для них. Кроме того, учитывалась возможность и сложность применения математического аппарата, с помощью которого находятся параметры моделей.

2. Анализ МНПС на предмет возможности их использования при учете фактора вторичных дефектов программных средств

Одним из параметров, с помощью которого проводится определение показателей надежности ПС, есть интенсивность проявления дефектов (функция риска). Исходя из этого, был проведен выбор моделей для анализа возможности их использования при учете фактора вторичных дефектов (табл. 1). Проанализированы основные допущения моделей: вид распределения между моментами проявления дефектов, интенсивность обнаружения дефектов, количество обнаруженных и устраненных дефектов на временном интервале тестирования, аналитическое выражение функции риска.

Для количественного анализа проявления вторичных дефектов, в первую очередь, проанализированы функции риска моделей.

Геометрические модели. Аналитическое выражение функции риска

$$\lambda(t) = DK^{i-1}, \quad (2)$$

где $D = R(0)$ – исходное значение функции риска, K – знаменатель геометрической прогрессии ($0 < K < 1$). Прибавление к данному выражению слагаемого n^{BH} , характеризующее

¹ Одарущенко О.Н. CASE-оценка критических программных систем...; Харченко В.С. Анализ сценариев и определение параметров для оценки надежности программных средств с учетом вторичных дефектов / В.С. Харченко, О.Н. Одарущенко, А.А. Руденко, Е.Б. Одарущенко // Системи управління, навігації та зв'язку. – 2011. – Випуск 2 (18). – С. 273-280.



го число вторичных дефектов, не соответствует размерности; прибавление слагаемого в степень числа K уменьшает значение функции риска, что не соответствует реалиям.

Модель Шнайдевинда. Аналитическое выражение функции риска

$$\lambda(t) = B \exp(-Ki), \quad (3)$$

где $B > 0$ и $K > 0$ – параметры модели, i – порядковый номер интервала.

Функцию риска можно модифицировать следующим образом:

$$\lambda(t) = B \exp(-K(i + n^{\text{BH}})) \quad (4)$$

или

$$\lambda(t) = B \exp(-Ki + n^{\text{BH}}). \quad (5)$$

Однако в модели предполагается разная степень влияния на процесс предсказания дефектов, в зависимости от времени.

Существуют три подхода к предсказанию дефектов.

Первый подход используется в случаях, когда данные всех интервалов необходимы для предсказания будущего состояния ПС, что отражает равнозначность данных из всех интервалов.

Второй – когда есть основания считать, что произошло некоторое существенное изменение в процессе обнаружения дефектов и для предсказания требуются только данные последних $m - s + 1$ интервалов. Это свидетельствует о том, что более ранние данные оказывают меньшее влияние (или не оказывают вообще) на прогнозирование надежности.

Третий подход является переходным между первыми двумя и отображает процесс выявления дефектов на более поздних стадиях.

Таким образом, учет рассмотренных выше подходов не позволяет получить однозначную оценку надежности ПС с учетом вторичных дефектов, что свидетельствует о нецелесообразности использования данной модели. Кроме того, нужно учитывать возможность обнаружения на интервале тестирования более одного дефекта, что недостаточно согласует величины i и n^{BH} .

Модель Дюзана. Общее число дефектов, обнаруженных к произвольному моменту времени, распределено по закону Пуассона со средним значением $m(t)$, где $m(t) = \alpha t^\beta$.

Отсюда следует, что

$$\frac{m(t)}{t} = \frac{\alpha t^\beta}{t}. \quad (6)$$

Числитель выражения (6) характеризует ожидаемое число дефектов за время t , а знаменатель – общее время тестирования t .

Величина n^{BH} трудно согласуется с параметром модели β . Для реализации этой модели требуется знание моментов проявления дефектов, то есть временная характеристика, что не согласуется с количественной характеристикой n^{BH} .

Таким образом, показана невозможность использования геометрических моделей, модели Шнайдевинда и модели Дюзана для оценки надежности ПС с учетом вторичных дефектов.

Сравнение функций риска и модифицированных функций риска других МНПС представлено в табл. 2. Эти модели имеют аналогичные аналитические выражения функций риска и содержат общие для всех параметры: K – коэффициент пропорциональности; B – исходное (неизвестное) число оставшихся в ПС дефектов.

Функции риска моделей, представленных в табл. 2, легко модифицируются прибавлением слагаемого n^{BH} , что дает основание допускать возможность их использования для оценки надежности с учетом вторичных дефектов.

Дальнейшее исследование предусматривает анализ допущений, принятых для моделей, а также, анализ непосредственно самих функций риска моделей Джелинского-Моранды, простой экспоненциальной модели, Шика-Уолвертона и Липова.

Анализ МНПС на предмет возможности их использования с учетом вторичных дефектов

Модель	Распределение между моментами обнаружения дефектов	Интенсивность обнаружения дефектов	Количество обнаруженных дефектов на интервале	Количество устраненных дефектов на интервале	Обнаружение дефектов	Аналитическое выражение функции риска (возможность модификации с учетом фактора вторичных дефектов)
Джелилинско-го-Моранды	const	пропорционально текущему числу дефектов	1	1	равновероятно	+
Простая экспоненциальная	экспоненциальное	пропорционально текущему числу дефектов	1	1	равновероятно	+
Шика-Уолвертона	пропорционально времени тестирования после обнаружения последнего дефекта	пропорционально текущему числу дефектов	1	1	равновероятно	+
Липова	Пуассона	постоянная на всем интервале тестирования	f_i	n_i	равновероятно	+
Геометрические	экспоненциальное	геометрическая прогрессия	1	1	неравновероятно	-
Шнайдевинда	неоднородный процесс Пуассона с экспоненциально уменьшающимся уровнем обнаружения дефектов	пропорционально текущему числу дефектов	n_i	n_i	неравновероятно	+
Дюэна	Пуассона	Зависит от моментов времени t_i проявления дефектов	1	1	равновероятно	-

Из табл. 1 следует, что по четырем позициям из пяти основные допущения моделей Джелилинского-Моранды, простой экспоненциальной модели, Шика-Уолвертона совпадают, за исключением вида распределения между моментами обнаружения дефектов, что в данном исследовании не является принципиальным. Преимущество данных моделей над моделью Липова состоит в сравнении следующих показателей:

- интенсивность обнаружения дефектов пропорциональна текущему числу дефектов против постоянной на всем интервале тестирования;
- количество обнаруженных (один) и устраненных дефектов (один) на интервале равно против f_i и n_i ($f_i \geq n_i$) соответственно.

Таким образом, модель Липова использовать менее целесообразно, чем три другие рассматриваемые модели.

Таблица 2



Модель	Функция риска	Модифицированная функция риска	Параметры функций риска
Джелинского-Моранды	$\lambda(t_i) = K(B - (i - 1))$	$\lambda(t_i) = K(B - (i - 1) + n^{\text{BH}})$	t – произвольная точка времени между обнаружением $i - 1$ и i -го дефекта
Простая экспоненциальная	$\lambda(t) = K(B - N(t))$	$\lambda(t) = K(B - N(t) + n^{\text{BH}})$	$N(t)$ – число обнаруженных к моменту времени t дефектов
Шика-Уолвертона	$\lambda(t_i) = K(B - (i - 1)) X_i$	$\lambda(t_i) = K(B - (i - 1) + n^{\text{BH}}) X_i$	X_i – время тестирования, прошедшее от момента t_{i-1} обнаружения $i - 1$ -го дефекта до текущего момента t_i
Липова	$\lambda(t) = K(B - F_{i-1}),$ $t_{i-1} \leq t \leq t_i$	$\lambda(t) = K(B - F_{i-1} + n^{\text{BH}})$	$F_{i-1} = \sum_{j=1}^{i-1} m_j$ – общее число скорректированных к моменту t_{i-1} дефектов, а t_i – время конца i -го интервала тестирования

Учитывая практически полное соответствие основных допущений моделей Джелинского-Моранды, простой экспоненциальной модели и Шика-Уолвертона, дальнейший анализ направлен на уровень сложности математического аппарата при их использовании. В табл. 2 (4 столбец) показаны преимущества модели Джелинского-Моранды перед другими моделями: функция риска модели Джелинского-Моранды не содержит кумулятивной характеристики обнаруженных дефектов (в отличие от простой экспоненциальной модели) и характеристики, определяющей время тестирования (в отличие от модели Шика-Уолвертона).

Выбор для исследования модели Джелинского-Моранды продиктован, в первую очередь, меньшей сложностью математического аппарата по сравнению с другими вышеупомянутыми моделями при других условиях (допущениях), не имеющих концептуальных преимуществ. Далее показана невозможность непосредственного нахождения распределения вторичных дефектов на основе модифицированных моделей, что не исключает вариант его нахождения при комплексировании по входным данным с моделями других классификационных признаков¹.

3. Прогнозирование проявления вторичных дефектов программных средств по статистическим данным обнаружения дефектов

Сложности аналитического определения числа или распределения вторичных дефектов обуславливают необходимость постановки и решения задачи прогнозирования их проявления по статистическим данным обнаружения дефектов, полученным в результате тестирования или эксплуатации ПС. Основной принцип прогнозирования вторичных дефектов состоит в нахождении флуктуаций эмпирического распределения данных обнаружения дефектов от теоретического, при этом, введением соответствующего коэффици-

¹ Руденко А.А. Указ. соч.

ента учитываются факторы, характеризующие этап тестирования или эксплуатации (вероятность проявления дефектов равно как и внесение вторичных дефектов более вероятно на раннем этапе)¹. Предполагается, что причиной отклонения эмпирических данных от теоретических являются вторичные дефекты, причем их количество определяется разностью модуля величины отклонения эмпирического значения от теоретического и среднего квадратического отклонения, умноженного на коэффициент $\frac{1}{n+1-x}$, где n – число интервалов тестирования (число обнаруженных дефектов), x – порядковый номер интервала (обнаруженного дефекта). Результат округляется до целых чисел.

Теоретическое распределение числа дефектов получаем на основе определения уравнения линии регрессии по эмпирическому распределению числа дефектов.

В зависимости от вида линии регрессии возможны некоторые нюансы учета (или неучета) вторичных дефектов. Рассмотрен случай линии регрессии вида $y = a + \frac{b}{x}$ ($a > 0, b > 0$)². Очевидно, что при любых $x > 0$ $y \geq a$, при увеличении x y асимптотически приближается к a . Поэтому на некоторых поздних интервалах тестирования (эксплуатации) возможно прогнозирование вторичного дефекта при теоретическом значении большем соответствующего эмпирического. В этих случаях учитывать вторичный дефект не нужно.

Использование корреляционного анализа позволяет получить другие, отличные от наиболее часто применяемых экспоненциального и распределения Пуассона, распределения, характеризующие зависимость числа дефектов от времени.

Исследования показали, что примерно в пяти процентах случаев при прогнозировании вторичных дефектов по статистическим данным могут возникать проблемы при эмпирическом определении линии регрессии или некоррелированности эмпирических и теоретических данных. Это объясняется неслучайными факторами внесения дефектов.

4. Определение показателей надежности ПС при учете фактора вторичных дефектов

Учет вторичных дефектов позволяет уточнить интенсивность проявления дефектов и использовать полученный параметр для нахождения функции готовности. Каждый прогнозируемый вторичный дефект увеличивает на один число внутренних фрагментов. Проведенные расчеты для значения функции готовности при разном числе внесенных дефектов и соответствующие расчеты средствами имитационного моделирования показали, что среднее отклонение по абсолютной величине одних результатов от других соответствует многофрагментной модели для случая, который отвечает спрогнозированному методами корреляционного анализа числу вторичных дефектов.

Определение показателей надежности ПС при учете фактора вторичных дефектов содержит взаимосвязанные этапы исследования: анализ и модификация МНПС, разработка сценариев внесения и устранения дефектов, выбор среди них базового сценария, который соответствует модифицированной МНПС, прогнозирование вторичных дефектов с использованием корреляционного анализа, уточнение величины интенсивности проявления дефектов, нахождение функции готовности средствами многофрагментного и имитационного моделирования, сравнение полученных результатов (рисунок).

Выводы

Результатом проведенных исследований является разработанный метод оценки надежности ПС с учетом вторичных дефектов, базирующегося на использовании модифицированной модели Джелинского-Моранды и процедуре прогнозирования вторичных дефектов по статистическим данным.

Усовершенствованы вероятностные МНПС на основе учета параметров вторичных дефектов путем модификации функций риска этих моделей, что позволяет повысить точность оценки показателей надежности.

¹ Харченко В.С. Указ. соч.

² Там же

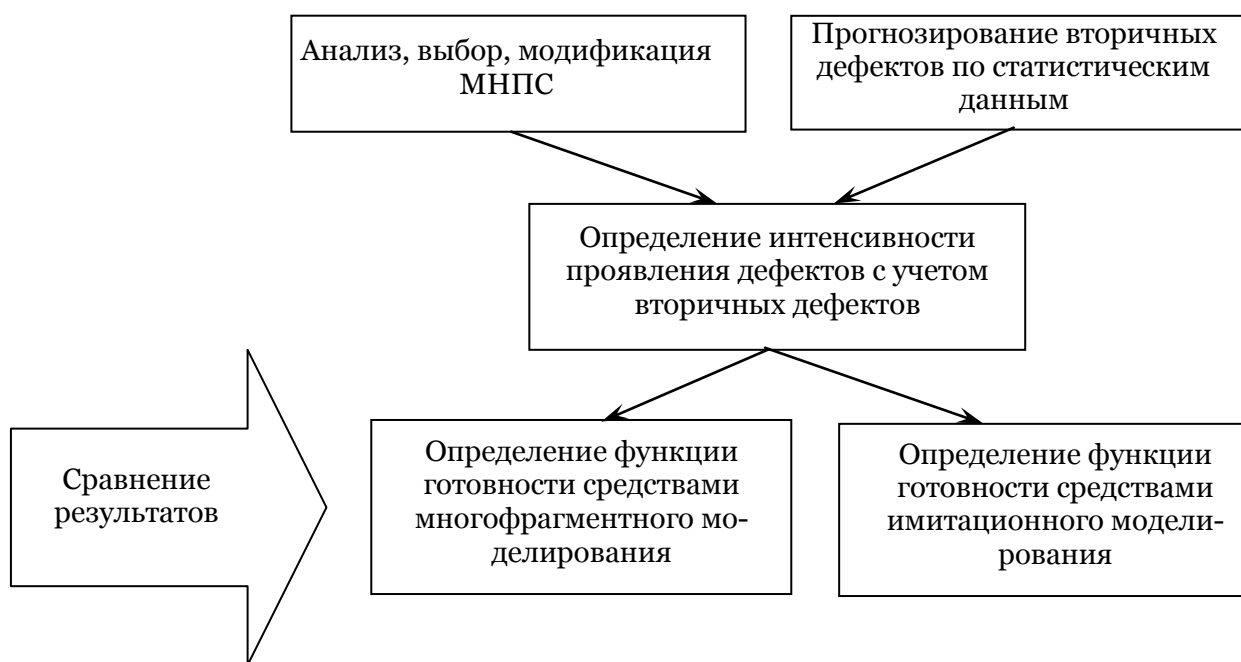


Рис. Последовательность оценки надежности ПС с учетом вторичных дефектов

Разработан метод прогнозирования вторичных дефектов по статистическим данным обнаружения дефектов на основе корреляционного анализа, что позволяет более точно определять параметры для оценки надежности ПС.

Данные методы апробированы при оценивании надежности ряда систем¹, в ходе которого была получена более высокая точность, определяемая отклонением полученных аналитических моделей от трендов дефектов ПС.

Дальнейшие исследования могут быть направлены на использование предложенных моделей для оценивания готовности систем с учетом процессов восстановления, а также, исследовании возможностей совместного использования предложенного метода и метода, основанного на теории динамики программных систем².

FACTORS OF SECONDARY FAULTS IN OF SOFTWARE RELIABILITY ASSESSING

V.S. KHARCHENKO¹
A. A. RUDENKO²
O.N. ODARUSHCHENKO³
E.B. ODARUSHCHENKO²

¹) National aerospace university,
 N.E. Zhukovskiy named after
 N.E. Zhukovskiy «KhAI», Kharkov

²) Poltava national technical university
 named after Yuri Kondratyuk

³) Research and production company
 «Radiy», Kirovograd

e-mail:
 v_s_kharchenko@ukr.net
 olekrudenko@yandex.ru
 skifs2005@mail.ru
 skifs2007@mail.ru

The paper considers the problem of accounting of the secondary software faults in the process of software reliability assessing. The existed models of software reliability are analysed in point of view of their use taking into account the secondary faults. The method of software reliability estimation taking into account the secondary faults is suggested.

Keywords: software, reliability, secondary software fault, failure rate, software reliability model

¹ Одарущенко О.Н. CASE-оценка критических программных систем...

² Маєвський Д.А. Структурна динаміка програмних систем...