

УДК 004.519.217

Д.А. Маевский, канд. техн. наук

МОДЕЛИРОВАНИЕ НАДЕЖНОСТИ В ТЕОРИИ ДИНАМИКИ ПРОГРАММНЫХ СИСТЕМ

Аннотация. На базе теории динамики программных систем разработаны теоретические основы и метод моделирования надежности. Проведен сравнительный анализ созданной модели надежности программного обеспечения с существующими. Показано, что созданная модель отличается от существующих значительно большей точностью моделирования для всех классов программных систем.

Ключевые слова: надежность, динамика программных систем, модели надежности.

D.A. Maevsky, PhD

RELIABILITY MODELLING IN DYNAMICS OF PROGRAMMING SYSTEMS THEORY

Abstract. In the article, on the base of theory of dynamics of the programming systems theoretical bases and method of design of reliability are worked out. The comparative analysis of the created model of reliability of software is conducted with existing. It is shown that the created model differs from existing considerably greater exactness of design for all classes of the programmatic systems.

Keywords: reliability, dynamics of the programming systems, model of reliability.

Д.А. Маевський, канд. техн. наук

МОДЕЛЮВАННЯ НАДІЙНОСТІ В ТЕОРІЇ ДИНАМІКИ ПРОГРАМНИХ СИСТЕМ

Анотація. На базі теорії динаміки програмних систем розроблено теоретичні основи та метод моделювання надійності. Проведено порівняльний аналіз створеної моделі надійності програмного забезпечення з існуючими. Показано, що створена модель відрізняється від існуючих значно більшою точністю моделювання для усіх класів програмних систем.

Ключові слова: надійність, динаміка програмних систем, моделі надійності.

Введение

Компьютерные информационные системы (ИС), широко используемые во всех областях человеческой деятельности, кроме несомненных выгод от своего применения несут и несомненные риски. Эти риски связаны с последствиями их отказов и сбоев, которые могут стать причиной крупных финансовых потерь, человеческих жертв, техногенных и экологических катастроф. Эти риски обусловлены, в основном дефектами в программном обеспечении (ПО) ИС.

Гарантировать отсутствие дефектов нельзя ни для одной ИС, потому актуальным является вопрос прогнозирования их количества и интенсивности выявления.

Для прогнозирования используются модели надежности программного обеспечения (МНПО) [8]. На сегодняшний день насчитывается около тридцати различных моделей и их модификаций, которые отличаются одна от другой подходами к процессу моделирования и начальными предположениями.

Такое обилие разнообразных моделей, описывающих один и тот же физический процесс, не может являться показателем хорошей изученности вопроса. Наоборот, это свидетельствует о непонимании нами процессов, которые лежат в основе проявления ошибок в ПО. Подтверждением этому может служить то, что ни одна из известных МНПО не обеспечивает приемлемого качества моделирования для широкого класса программных систем [4].

Как показывают приведенные в статье исследования, точность моделирования существующих моделей существенно зависит от исследуемой программной системы и показывает существенный непредсказуемый разброс значений. Поэтому актуальным является разработка универсальной МНПО, позволяющей моделировать надежность любых программных систем (ПС) на всех стадиях их разработки.

1. Основы теории динамики программных систем

В работах [1, 2] автором предложена теория динамики программных систем

© Маевский Д.А., 2011

(ДПС), основанная на концепции потоков дефектов, направленных из программной системы в окружающую среду (выявление и исправление дефектов) и из среды в программную систему (вторичные дефекты, вносимые при исправлении выявленных). В теории ДПС программная система рассматривается как открытая неравновесная система и предполагается, что она подчиняется общим законам динамики неравновесных систем. На основании этого математической моделью процесса выявления дефектов является автономная система дифференциальных уравнений

$$\begin{cases} \frac{df_1}{dt} = -A_1 f_1 - A_2 f_2 \\ \frac{df_2}{dt} = -A_2 f_1 - A_1 f_2 \end{cases}, \quad (1)$$

где f_1 – количество дефектов в системе, образующих выходной поток, f_2 – количество дефектов, внесенных в систему входным потоком, A_1 и A_2 – коэффициенты влияния, имеющие размерность время⁻¹.

Физический смысл коэффициента A_1 в обоих уравнениях одинаковый – это коэффициент пропорциональности между количеством дефектов и соответствующим этим дефектам потоком. Коэффициент A_2 характеризует взаимное влияние входного потока дефектов на выходной поток и наоборот, выходного потока на входной.

Производные $\frac{df_1}{dt}$ и $\frac{df_2}{dt}$ в (1) по своему физическому смыслу являются интенсивностью отказов, которая является основным параметром, позволяющим рассчитать показатели надежности и прогнозировать ее изменения во время эксплуатации системы [3]:

$$\lambda_1(t) = \left| \frac{df_1}{dt} \right|, \quad \lambda_2(t) = \left| \frac{df_2}{dt} \right|.$$

Здесь производные взяты по модулю, так как интенсивность отказов всегда положительна, а в случаях, когда f_1 или f_2 уменьшаются, их производные будут отрицательными.

Решения системы (1) при начальных условиях $f_1(0) = F_{10}$, $f_2(0) = 0$ являются зависимостями:

для количества дефектов, обуславливающих выходной поток:

$$f_1(t) = F_{10} e^{-A_1 t} \operatorname{ch}(A_2 t),$$

для интенсивности отказов, обусловленной выходным потоком:

$$\lambda_1(t) = A_1 F_{10} e^{-A_1 t} \operatorname{ch}(A_2 t) + A_2 F_{10} e^{-A_1 t} \operatorname{sh}(A_2 t),$$

для количества дефектов, внесенных в ПС входным потоком:

$$f_2(t) = F_{10} e^{-A_1 t} \operatorname{sh}(A_2 t).$$

для интенсивности отказов, обусловленной входным потоком:

$$\lambda_2(t) = A_1 F_{10} e^{-A_1 t} \operatorname{sh}(A_2 t) + A_2 F_{10} e^{-A_1 t} \operatorname{ch}(A_2 t).$$

В этих выражениях начальное количество дефектов в ПС в момент времени $t = 0$ обозначено F_{10} . На основании приведенных выражений можно построить модель надежности ПО.

2. Разработка модели надежности ПО

Для построения МНПО необходимо определиться с набором входных данных, на основании которых будет строиться модель, записать математические зависимости, определяющие показатели надежности, и привести методику определения коэффициентов в этих зависимостях исходя из экспериментально определенных входных данных.

В качестве входных данных в разрабатываемой модели принят временной ряд, задающий кумулятивное число выявленных дефектов. Как отмечается в [4], несмотря на то, что дефекты образуют потоки с четко определенными закономерностями, процесс их выявления имеет значительную неопределенность. Поэтому учет при моделировании каждого отдельного дефекта усложняет анализ результатов за счет имеющего место значительного «шума». В связи с этим МНПЗ, входными данными которых являются моменты выявления каждого дефекта, не могут обеспечить точность моделирования по причине того, что неточными уже являются их входные данные. Временной ряд, образующий кумулятивное количество дефектов, является более точным, потому что случайный учет или не учет каждого конкретного дефекта не может отразиться на общей тенденции изменения этого ряда. Фактически временной ряд, образованный кумулятивным количеством дефектов, уже освобожден от случайной составляющей и является трендом.

Итак, для построения модели, математические зависимости, полученные из теории ДПС, должны быть трансформированы для операций с кумулятивными трендами дефектов. С этой целью используем зависимости для интенсивности отказов выходного потока $\lambda_1(t)$ и входного потока $\lambda_2(t)$, приведенные выше.

Согласно теории надежности ПО [4], зависимость между кумулятивным количеством дефектов $\mu(t)$ и интенсивностью $\lambda(t)$ задается формулой

$$\mu(t) = \int_0^t \lambda(t) dt.$$

Для кумулятивного тренда исходного потока получим выражение:

$$\begin{aligned} \mu_1(t) = & \int_0^t \frac{F_{10}}{2} \left[(A_1 + A_2) e^{(A_2 - A_1)t} \right] + \\ & + \int_0^t \frac{F_{10}}{2} \left[(A_1 - A_2) e^{-(A_2 + A_1)t} \right] dt. \end{aligned} \quad (2)$$

Не трудно видеть, что при $A_1 = A_2$ первообразная от $\lambda_1(t)$ не существует, потому что разность между A_1 и A_2 будет равняться нулю. Поэтому нахождение кумулятивного количества дефектов исходного потока рассмотрим отдельно для двух случаев.

Случай 1. $A_1 \neq A_2$.

В этих условиях первообразная для (2) существует всегда, потому после интегрирования получим

$$\begin{aligned} \mu_1(t) = & \frac{F_{10}}{2} \left[\left(\frac{A_2 + A_1}{A_2 - A_1} e^{A_2 t} + \frac{A_2 - A_1}{A_2 + A_1} e^{-A_2 t} \right) e^{-A_1 t} \right] - \\ & - F_{10} \frac{A_1^2 + A_2^2}{A_2^2 - A_1^2}. \end{aligned} \quad (3)$$

Случай 2. $A_1 = A_2$.

Перед нахождением первообразной преобразуем выражение (2), учитывая, что $A_1 = A_2$. Получим: $\lambda_1(t) = F_0 A_1$, откуда

$$\mu_1(t) = F_{10} A_1 \int_0^t dt = F_{10} A_1 t. \quad (4)$$

Выражение (4) при $A_1 = A_2$ хорошо соотносится с ожидаемым результатом. Действительно, при этом на каждый дефект, который выносится из ПС, приходится один вторичный дефект, который вносится в нее. Поэтому общее количество дефектов в системе ос-

тается неизменным и частота вынесения дефектов благодаря этому тоже неизменная. Итак, при $A_1 = A_2$ ожидается линейная зависимость кумулятивного количества вынесенных дефектов от времени, которая и получена из теории динамики программных систем.

Для кумулятивного тренда входного потока (вторичных ошибок) из уравнения для $\lambda_2(t)$ после интегрирования получим

$$\begin{aligned} \mu_2(t) = & \frac{F_{10}}{2} \left(\frac{A_2 + A_1}{A_2 - A_1} e^{A_2 t} - \frac{A_2 - A_1}{A_2 + A_1} e^{-A_2 t} \right) e^{-A_1 t} - \\ & - 2F_{10} \frac{A_1 \cdot A_2}{A_2^2 - A_1^2} \end{aligned} \quad (5)$$

при $A_1 \neq A_2$ и

$$\mu_2(t) = F_{10} A_1 \cdot t \quad (6)$$

при $A_1 = A_2$.

Сравнивая выражения (4) и (6), можно заметить, что при $A_1 = A_2$ кумулятивные тренды выходного и входных потоков совпадают. Этот результат тоже соответствует физическим представлениям. При $A_1 = A_2$ количество дефектов, которое вносится в систему, равняется количеству выносимых. Из этого и следует, что их кумулятивные тренды тоже должны быть одинаковыми.

Таким образом, математической моделью надежности ПО являются выражения (3), (4) и (5), (6) соответственно для выходного и входного потока. Для практического применения модели необходимо разработать методику вычисления параметров модели по экспериментальным данным.

3. Методика экспериментального определения параметров модели

При экспериментальном определении параметров модели входными данными является кумулятивный тренд дефектов, выявленных на известном временном промежутке. Определению подлежат параметры модели – коэффициенты влияния A_1 и A_2 , а также начальное количество дефектов в системе F_{10} .

Определение параметров будем производить в два этапа. На первом выполняется предварительная оценка параметров, а на втором – их уточнение.

Для предварительной оценки считаем, что входной поток дефектов отсутствует, по-

этому коэффициент $A_2 = 0$. В этом случае кумулятивная зависимость выявленных дефектов представляет собой экспоненту

$$\mu_1(t) = F_{10} - F_{10}e^{-A_1 t}. \quad (7)$$

В этой зависимости величина, обратная коэффициенту A_1 , – постоянная времени процесса $\tau = \frac{1}{|A_1|}$, подкасающаяся к экспоненте. В свою очередь величина подкасающейся может быть определена как показано на рис. 1.

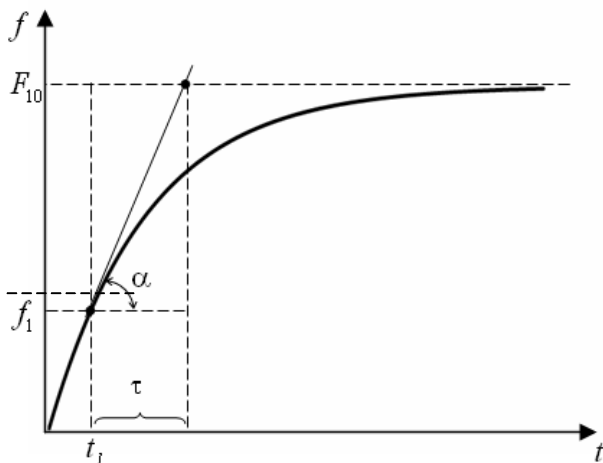


Рис. 1. К определению коэффициента A_1

Из рис. 1 следует, что величина подкасающейся может быть выражена как

$$\tau = \frac{F_{10} - f_1}{\text{tg}\alpha},$$

где угловой коэффициент $\text{tg}\alpha$ определяется из формулы

$$\text{tg}\alpha \approx \frac{f_1 - f_2}{t_1 - t_2}.$$

Для повышения точности расчетов следует определять угловой коэффициент $\text{tg}\alpha$ для каждой двух последовательных точек экспериментальной кумулятивной кривой. В качестве неизвестной величины F_{10} можно использовать последнюю точку кумулятивной кривой f_n . Определенные таким образом для каждой последовательной пары точек значения $\tau = \frac{1}{A_1}$ следует усреднить. Из найденного среднего значения τ находим приближение для коэффициента влияния A_1

$$A_1 = \frac{1}{\tau}. \quad (8)$$

Приближенное значение для F_{10} может быть получено из выражения:

$$F_{10} = \frac{f_i}{1 - e^{-A_1 t_i}}. \quad (9)$$

Для повышения точности найденные таким образом значения F_{10} усредняем по всем точкам экспериментальной кумулятивной кривой.

Следует учесть, что, полагая коэффициент $A_2 = 0$, то есть исключая влияние входного потока, мы получаем завышенные оценки для F_{10} и A_1 .

После определения приближений для F_{10} и A_1 следует итерационно уточнить эти значения. Необходимо иметь в виду зависимость $\mu_1(t)$ от коэффициентов A_1 и A_2 . Поверхность $\mu_1(t)$ в зависимости от A_1 и A_2 показана на рис. 2.

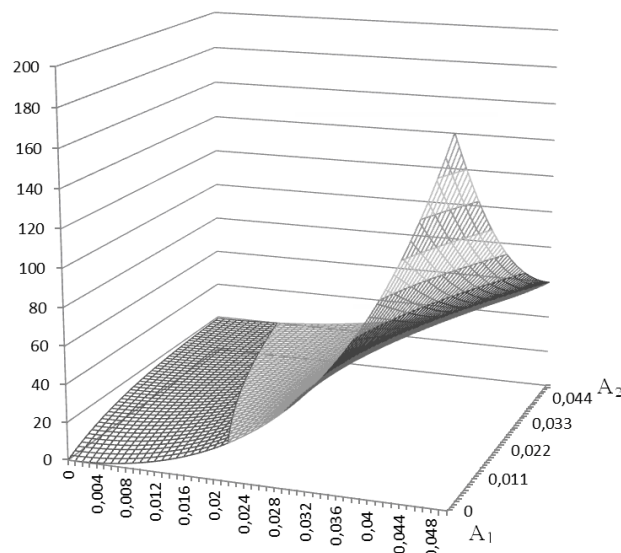


Рис. 2. Зависимость $\mu_1(t)$ от коэффициентов влияния

Из рис. 2 следует, что с изменением коэффициента A_2 значения $\mu_1(t)$ меняются незначительно. Поэтому уточнение параметров модели следует начать именно с коэффициента A_2 : допущенные здесь ошибки незначительно скажутся на точности процесса. Для оценки точности определения параметров используем критерий среднего квадратичного отклонения (СКО):

$$\text{СКО} = \frac{\sum_{i=1}^n (f_{i0} - f_{ic})^2}{n}, \quad (14)$$

где n – количество точек в экспериментальной кумулятивной кривой; f_{io} – а наблюдаемое зафиксированное (observed) значение i -й точки кривой; f_{ic} – вычисленное (calculated) с данными параметрами значение.

Для уточнения параметров модели предложен следующий алгоритм:

Шаг 1. Определить начальное значение SKO_b при полученных по формулам (8) и (9) приближениях для A_1 и F_{10} . При этом положить $A_2 = 0$.

Шаг 2. Изменяя коэффициент с шагом $\Delta A_2 = \frac{1.5A_1}{10}$ пока текущее значение SKO_x меньше, чем SKO_b , получить уточненное значение для A_2 . Положить $SKO_b = SKO_x$.

Шаг 3. Изменяя значение F_{10} в диапазоне от 0,5 до 1,5 от полученного в шаге 1 приближения пока текущее значение SKO_x меньше, чем SKO_b , найти уточненное значение для F_{10} . Положить $SKO_b = SKO_x$.

Шаг 4. Изменяя коэффициент A_1 в диапазоне от 0,5 до 1,5 от полученного в шаге 1 приближения пока текущее значение SKO_x меньше, чем SKO_b , найти уточненное значение для A_1 . Положить $SKO_b = SKO_x$.

Шаг 5. Повторить шаг 2 – шаг 4 для следующего этапа уточнения. Процесс уточнения считается законченным, если на очередном этапе достигнуто значение $SKO_b \leq \varepsilon$, где ε – заданная точность расчета.

4. Проверка разработанной модели надежности

Для оценки точности моделирования по разработанной МНПО и сравнения ее с существующими моделями были использованы данные по выявленным дефектам в двадцати разных программных системах [8, 5]. Для увеличения точности моделирования каждый ряд наблюдений разбивался на интервалы, в течение которых закон изменения во времени кумулятивных кривых выявленных дефектов оставался неизменным. Всего обработано 123 ряда наблюдений. Для сравнения точности моделирования, наряду с описанной МНПО, были взяты известные модели надежности, охватывающие все существующие классы моделей. В их основе лежат разные концепции, что дает возможность объ-

ективно оценить точность и сравнить модели между собой. В качестве исследуемых взяты модели Джелинского-Моранды [9], неоднородного пуассоновского процесса [7], Шнайдервинда [14], Мусы [11], Вейбулла [13], S-образная модель [15], модель Дюзна [6], геометрическая модель Моранды [10] и логарифмическая модель Мусы-Окумото [12]. Всего выполнено 1230 оценок моделирования надежности. Результаты моделирования приведены на рис. 3.

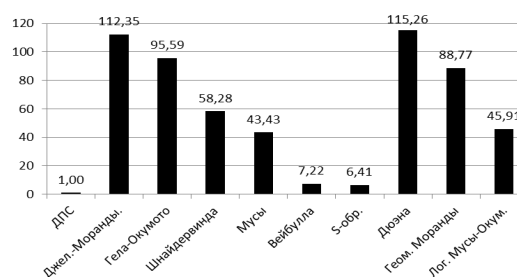


Рис. 3. Результаты сравнения МНПО

В диаграмме, приведенной на рис. 3, показаны значения SKO, полученные в результате сравнения. Для удобства SKO предлагаемой МНПО на основе ДПС принято равным единице. Из диаграммы видно, что для всех классов исследуемого ПО модель ДПС показала результат по точности более, чем в шесть раз превосходящий результат наилучшей из известных МНПО – S-образной модели.

Выводы

Проведенный сравнительный анализ разных МНПО показал, что предложенная в работе модель на основе теории ДПС стабильно показывает наилучшие результаты в сравнении с другими известными моделями.

Стабильность результатов моделирования является существенным фактором, позволяющим избежать этапа выбора модели надежности для каждого отдельного программного продукта. Именно этап выбора, как показано в [2], является самым трудоемким и математически неопределенным.

Таким образом, полученные результаты являются практическим подтверждением теории ДПС, а разработанная на их базе модель надежности может быть использована для моделирования и прогнозирования показателей надежности широкого класса программных систем.

Список использованной литературы

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

2. Маевский Д.А. Динамика программных систем и модели их надежности. // Сб. *Радиоэлектронные и компьютерные системы*. – 2011. – №.2. – С. 45-54.

3. Смагин В.А. Основы теории надёжности программного обеспечения: учеб. пособие / В.А. Смагин. – СПб.: ВКА им.А.Ф. Можайского, 2008. – 331 с.

4. Харченко В.С. Методы моделирования и оценки качества и надежности программного обеспечения / В.С. Харченко, В.В. Скляр, О.М. Тарасюк. Учеб. пособие. – Харьков: Нац. аэрокосм. ун-т ХАИ, 2004. – 159 с.

5. Android – An Open Handset Alliance Project [Электронный ресурс]. Режим доступа: <http://code.google.com/p/android/issues/list>

6. Duan J.T. Lurning Curve Approach to Reliability Monitoring // *IEEE Trans. on Aerospace*. – 1964. – Vol. 2. – P. 563-566.

7. Goel, A.L., Time-Dependent Error-Detection Rate Model for Software and Other Performance Measures / A.L. Goel, K. Okumoto // *IEEE Transactions on Reliability*, v. R-28, № 5, August. – 1979. – С. 206–211.

8. Lyu M.R. Handbook of Software Reliability Engineering, McGraw-Hill Company, 1996. – 805 p.

9. Moranda P.B. Final Report of Software Reliability Study. – / P.B. Moranda, J. Jelinski // McDonnell Douglas Astronautics Company. MDC Report № 63921, dec. – 1972. – 51 с.

10. Moranda, P.B. Event-Altered Rate Models for General Reliability Analysis / P.B. Moranda // *IEEE Transactions on Reliability*. Vol.R-28. – No. 5. – 1979. – С.376-381

11. Musa J.D. Validity of Execution time theory of software reliability // *IEEE Trans. on Reliability*. – 1979. – № 3.– С.199–205.

12. Musa J.D. Okumoto K. A Logarithmic Poisson Time Model for Software Reliability Measurement /Proc. Sevent International Conference on Software Engineering. – Orlando, Florida: – 1984. – P. 230-238.

13. Quadri, S. M. K. Software Reliability Growth Modeling with New Modified Weibull Testing–effort and Optimal Release Policy / S. M. K. Quadri, N. Ahmad // *International Journal of Computer Applications*. – Vol. 6. – 2010. –№ 12. – С. 1–10.

14. Schneidewind, N.F. Software Reliability Model with Optimal Selection of Failure Data / N.F. Schneidewind // *IEEE Transactions on Software Engineering*. – Vol. 19. – No. 11. Nov. – 1993. – P. 1095-1104.

15. Yamada, S. S-Shaped Reliability Growth Modeling for Software Error Detection / S. Yamada, M. Ohba, S. Osaki // *IEEE Transactions on Reliability*. Vol. R-32. No. 5, Dec. – 1983. – P. 475-478.

Получено 10.10.2011

References

1. Maevsky D.A. Structural dynamics of programming systems and prognostication of their reliability at presence of secondary defects. / *Proc. Radio-electronic and computer systems*. - 2010. – № 3 – P. 103–109 [in Ukrainian].

2. Maevsky D. Dynamics of programming systems and their reliability models. / D.A. Maevsky / *Proc. Radio-electronic and computer systems*. – 2011. – № 2. – P. 45-54/ [in Russian].

3. Smagin V.A. Fundamentals of the Theory of software reliability: Textbook. Manual – St.Petersburg: PCA them.A.F. Mozhaisky, 2008. – 331 p.[in Russian].

4. Kharchenko V.S., Sklar V.V., Tarasyuk O.M. Methods for modeling and evaluating the quality and reliability of the programming security / Textbook.allowance – Kharkov: Nat. Aerospace. Univ"HAИ", 2004. – 159 [in Russian].

5. Android – An Open Handset Alliance Project : <http://code.google.com/p/android/issues/list> [in Russian].

6. Duan J.T. Lurning Curve Approach to Reliability Monitoring // *IEEE Trans. on Aerospace*. 1964. – Vol. 2. – P. 563-566 [in English].

7. Goel A.L., Okumoto K. Time-Dependent Error-Detection Rate Model for Software and Other Performance Measures / *IEEE Transactions on Reliability*. –V. R-28, № 5, August. – 1979. – P. 206-211 [in English].

8. Lyu M.R. Handbook of Software Reliability Engineering, McGraw-Hill Company, 1996. – 805 p. [in English].

9. Moranda P.B., Jelinski J. Final Report of Software Reliability Study. – /McDonnell Douglas Astronautics Company. MDC Report № 63921. Dec. – 1972. – 51 p. [in English].

10. Moranda P.B. Event-Altered Rate Models for General Reliability Analysis / IEEE Transactions on Reliability. Vol.R-28. – No. 5. – 1979. – P.376–381.

11. Musa J.D. Validity of Execution time theory of software reliability // IEEE Trans. on reliability. – 1979. – № 3.–P.199-205 [in English].

12. Musa J.D. Okumoto K. A Logarithmic Poisson Time Model for Software Reliability Measurement //Proc. Sevent International Conference on Software Engineering. – Orlando, Florida: – 1984. – P. 230–238 [in English].

13. Quadri, S. M. K., Ahmad N. Software Reliability Growth Modeling with New Modified Weibull Testing–effort and Optimal Re-

lease Policy / International Journal of Computer Applications Vol. 6. – 2010. – № 12. –P. 1–10 [in English].

14. Schneidewind, N.F. Software Reliability Model with Optimal Selection of Failure Data / N.F. Schneidewind //IEEE Transactions on Software Engineering. – Vol. 19. – No. 11. Nov. – 1993. – P. 1095–1104 [in English].

15. Yamada, S., Ohba M., Osaki S. Shaped Reliability Growth Modeling for Software Error Detection /IEEE Transactions on Reliability. – Vol. R-32. – No. 5, Dec. –1983. – P. 475–478 [in English].



Маевский

Дмитрий Андреевич,
канд. техн. наук, доцент,
зав. каф. теоретических ос-
нов и общей электротехники
Одесск. нац. политехн. ун-
та, тел. (8-048) 734-84-54