Національний університет “Львівська політехніка”

Сенів Максим Михайлович

УДК 004.052.3

ЗАСОБИ ПРОГНОЗУВАННЯ НАДІЙНОСТІ ПРОГРАМНОГО ПРОЕКТУ ІЗ ВРАХУВАННЯМ ПОКАЗНИКА ЙОГО СКЛАДНОСТІ

Спеціальність 01.05.03 – математичне та програмне забезпечення обчислювальних машин і систем

Автореферат дисертації на здобуття наукового ступеня
кандидата технічних наук

Львів – 2011

Дисертацією є рукопис.

Робота виконана в Національному університеті “Львівська політехніка” Міністерства освіти і науки, молоді та спорту України

Науковий керівник: доктор технічних наук, професор
Федасюк Дмитро Васильович,
Національний університет “Львівська політехніка”,
проректор з науково-педагогічної роботи,
завідувач кафедри програмного забезпечення

Офіційні опоненти: доктор технічних наук, доцент
Пелещишин Андрій Миколайович,
Національний університет “Львівська політехніка”,
професор кафедри інформаційних систем і мереж

кандидат технічних наук, доцент
Харченко Олександр Григорович,
Національний авіаційний університет (м. Київ),
професор кафедри комп’ютерних інформаційних технологій.

Захист відбудеться 9 червня 2011 р. о 14 годині на засіданні спеціалізованої вченої ради Д 35.052.05 у Національному університеті “Львівська політехніка” (79013, м. Львів, вул. С. Бандери, 12).

З дисертацією можна ознайомитися в науково-технічній бібліотеці Національного університету “Львівська політехніка” (79013, м. Львів, вул. Професорська, 1).

Автореферат розіслано ___ травня 2011 р.

Вчений секретар
спеціалізованої вченої ради,
доктор технічних наук, професор Р. А. Бунь

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Розвиток інформаційних технологій та поширення їх у всіх сферах людської життєдіяльності ставить щоразу складніші завдання перед розробниками програмного забезпечення. Програмні продукти стають дедалі складнішими, багатокомпонентними і вимагають спеціалізованого підходу для їх розроблення.

Отже, підвищуються і вимоги до надійності програм, виникає потреба у скороченні затрат на тестування та, відповідно, у прогнозуванні надійності розроблюваного програмного забезпечення (ПЗ).

В багатьох дослідженнях поняття надійності ПЗ виділяють із загальної теорії надійності технічних систем в окремий напрям. При застосуванні понять надійності до програмних засобів необхідно враховувати особливості і відмінності цих об'єктів від традиційних технічних систем, для яких спершу розроблялася теорія надійності. Принципова відмінність програмного забезпечення від апаратного полягає у динаміці процесу, а саме в тому, що програма за період використання не зношується, а навпаки, з плином часу виявляються і виправляються помилки, які не були знайдені раніше, тому ПЗ з часом вдосконалюється і покращується.

Для розв’язання задач оцінювання та прогнозування надійності ПЗ в даний час використовуються відповідні до ситуації моделі надійності. Модель надійності ПЗ – це математична модель, побудована для визначення залежності характеристик надійності ПЗ від ряду параметрів. Моделі можна класифікувати за різними ознаками. Зараз широко використовують тип моделей, які описують характеристики ПЗ на основі кількості виявлених помилок у програмному проекті.

В основу значної кількості моделей надійності ПЗ, предметом дослідження яких є кількість помилок у визначеному часовому інтервалі, покладено розподіл Пуассона (найпоширеніші моделі цього класу – Goel-Okumoto та S-подібна). Використання цього розподілу випадкових величин добре зарекомендувало себе в багатьох областях, де основна особливість полягає в обчисленні кількості незалежних подій протягом інтервалів часу.

Але в моделях на основі кількості виявлених помилок або зовсім не враховується такий показник, як складність програмного проекту, або враховується тільки в часткових випадках, зокрема має апріорно встановлені значення або не є параметром моделі. Саме це і обґрунтовує актуальність наукової задачі розроблення засобів для прогнозування надійності ПЗ з використанням показника складності програмного проекту.

Зв’язок роботи з науковими програмами, планами, темами. Дослідження, що становлять матеріал дисертації, безпосередньо пов’язані з науковим напрямком кафедри програмного забезпечення Національного

університету „Львівська політехніка”. Результати, викладені в дисертаційній роботі, отримані в межах таких держбюджетних науково-дослідних робіт:

- ДБ/ПНРЛ – „Розроблення моделей, методів та алгоритмів для автоматизованої оцінки показників надійності радіоелектронних та електромеханічних пристроїв і систем” (номер державної реєстрації 0110U001098), яка виконувалася у 2010 році. Дисертантом розроблено модель з динамічним показником складності проекту, здійснено формалізацію критерію достатності процесу тестування програмного забезпечення, розроблено метод оцінювання та прогнозування надійності програмного забезпечення на основі цієї моделі.
- ДБ/Діаграма – „Розробка методів та засобів розподілення обчислень в задачах теплового проектування електронних пристроїв нового покоління” (номер державної реєстрації 0108U000331), яка виконувалася в період з 2008 до 2009 року. Дисертантом розроблено інформаційну модель представлення даних середовища оцінювання надійності програмних систем та програмний модуль оцінювання та прогнозування надійності.
- ДБ/НМ – „Математичне та програмне забезпечення мережевих систем теплового проектування сучасних мікроелектронних пристроїв” (номер державної реєстрації 0106U000260), яка виконувалася в період з 2006 до 2007 року. Дисертантом проведено дослідження основних характеристик найбільш поширених криптоалгоритмів.

Мета і завдання дослідження. Метою дисертаційної роботи є створення та удосконалення засобів для прогнозування надійності програмного забезпечення з використанням динамічного показника складності програмного проекту.

Для досягнення цієї мети в дисертації сформульовано і розв’язано такі завдання:

1. Аналіз характеристик і параметрів надійності ПЗ та засобів її прогнозування.
2. Дослідження та удосконалення моделей оцінювання і прогнозування надійності ПЗ на основі кількості виявлених помилок з урахуванням рівня складності програмного проекту.
3. Формалізація критерію достатності процесу тестування ПЗ та дослідження поведінки показника складності програмного проекту.
4. Розроблення методу оцінювання та прогнозування надійності ПЗ на основі моделі надійності ПЗ з урахуванням показника складності проекту.
5. Розроблення інформаційного та програмного забезпечення на основі отриманих теоретичних результатів.
6. Верифікація розробленого математичного та програмного забезпечення, аналіз отриманих результатів.

Об'єктом дослідження є процес прогнозування надійності програмного забезпечення.

Предметом дослідження є моделі, методи та критерії оцінювання і прогнозування надійності програмного забезпечення.

Методи дослідження. При проведенні досліджень використано засоби і методи теорії ймовірностей і математичної статистики – для аналізу, розроблення та вдосконалення моделей оцінювання та прогнозування надійності ПЗ; для визначення критеріїв та точкових оцінок параметрів моделі – метод максимальної правдоподібності; для розв'язування системи трансцендентних рівнянь – модифікований метод Ньютона; для розроблення програмного забезпечення – методи об'єктно-орієнтованого і прикладного програмування. Для опрацювання експериментальних даних застосовано програмні засоби обчислення (пакети MathCad та Statistica).

Наукова новизна одержаних результатів:

1. Удосконалено моделі надійності ПЗ, де показник складності програмного проекту вперше введено параметром моделі, а форми кривих функції кумулятивної кількості несправностей та інтенсивності появи помилок більш адекватно відображають практичні результати тестування ПЗ.
2. Вперше формалізовано критерій достатності процесу тестування ПЗ, який базується на залежності динамічного показника складності проекту від кількості циклів тестування програмного продукту, що дає можливість зменшити похибку прогнозування загальної кількості помилок.
3. Вперше розроблено метод оцінювання та прогнозування надійності ПЗ на основі удосконаленої моделі надійності з динамічним показником складності проекту, який дає можливість обґрунтовано приймати рішення щодо продовження процесу тестування.
4. Удосконалено метод розрахунку кількісних характеристик надійності ПЗ шляхом врахування значення критерію достатності процесу тестування, що дає змогу знизити похибку при визначенні кількості залишкових несправностей.

Практичне значення одержаних результатів:

1. Розроблена модель надійності ПЗ придатна до використання на підприємстві зі створення ПЗ для оцінювання загальної кількості помилок в ПЗ та прогнозування частоти їх появи.
2. Розроблений алгоритм розрахунку числових показників надійності ПЗ дає змогу розрахувати такі параметри надійності ПЗ, як час напрацювання на відмову, загальну кількість помилок, що є придатним для числової оцінки надійності розроблюваного ПЗ на виробництві.

3. Значення критерію достатності процесу тестування ПЗ дає практичні рекомендації керівникам тестових підрозділів ІТ-компаній стосовно кількісної оцінки тривалості етапу тестування ПЗ.
4. Динамічний показник складності проекту є підставою для практичних рекомендацій керівникам проекту стосовно розподілу виробничих ресурсів між етапами життєвого циклу проекту.
5. Програмна система оцінювання та прогнозування надійності програмного забезпечення дає змогу отримати рекомендації щодо достатності чи потреби в продовженні процесу тестування та обчислити кількісні характеристики надійності ПЗ.

Результати дисертаційного дослідження впроваджені при розробці програмних продуктів для оцінювання та прогнозування їх надійності на підприємствах ТзОВ „СплайнСофвер”, ТзОВ „Елекс”, а також використані у навчальному процесі Національного університету “Львівська політехніка” при читанні курсу лекцій з дисципліни “Якість програмного забезпечення та тестування”, що підтверджено відповідними актами.

Особистий внесок здобувача. Усі наукові результати, представлені у дисертаційній роботі, отримані автором особисто. У друкованих працях, опублікованих у співавторстві, автору належать: розробка моделі з динамічним показником складності проекту [9, 11], порівняльний аналіз існуючих та розробленої моделей на реальних тестових прикладах, формалізація критерію достатності процесу тестування програмного забезпечення [1], проведення аналізу використання моделі надійності програмного забезпечення з динамічним показником складності проекту протягом життєвого циклу ПЗ [2], розробка методу оцінювання та прогнозування надійності програмного забезпечення на основі моделі з динамічним показником величини проекту [3], розробка інформаційної моделі представлення даних середовища оцінювання надійності програмних систем [12], розробка архітектури, формату представлення даних і реалізація програмної системи для оцінювання та прогнозування надійності програмного забезпечення [5, 10], розробка програмного модуля оцінювання та прогнозування надійності [4, 18], дослідження основних характеристик найбільш поширених криптоалгоритмів (DES, RSA, RC5) [6 – 8, 16, 17], розробка модуля захисту даних [13,14], розробка програмного засобу для автоматизації оцінювання надійності програмної системи теплового проектування [15].

Апробація результатів дисертації. Основні результати наукових досліджень доповідалися на: IX, X International Conferences „The Experience of Designing and Application of CAD Systems in Microelectronics „CADSM” (м.Поляна 2007, 2009); 2-ій Міжнародній конференції молодих науковців „CSE-2007” (м.Львів 2007); IV International Conference of Young Scientists MEMSTECH 2008 (м.Поляна 2008); 6-ій міжнародній конференції "Інтернет – Освіта – Наука

– 2008" ІОН-2008 (м. Вінниця); 4-ій та 5-ій Міжнародних науково-технічних конференціях "Комп'ютерні науки та інформаційні технології" CSIT (м. Львів 2009, 2010); X International Conference “Modern problems of radio engineering, telecommunications, and computer science” (м.Славське 2010).

Публікації. За результатами проведених досліджень опубліковано 18 друкованих праць, з яких 8 праць у наукових фахових виданнях з переліку, затвердженого ВАК України, одна стаття в іншому виданні та 9 тез доповідей у матеріалах міжнародних науково-технічних конференцій.

Структура та обсяг роботи. Дисертаційна робота складається із вступу, 4 розділів, висновків, списку використаних джерел і додатків. Основний зміст викладено на 135 сторінках друкованого тексту. Дисертація містить 38 рисунків, 12 таблиць, 123 літературних джерела, 3 додатки. Повний обсяг дисертації 173 сторінки.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету роботи та задачі дослідження, викладено наукову новизну та практичну цінність одержаних результатів, а також подано інформацію про особистий внесок здобувача, апробацію результатів роботи та наявні публікації.

В **першому розділі** висвітлено поняття надійності, критерії та показники надійності, основи надійності ПЗ, а також здійснено огляд та порівняння основних відомих моделей оцінювання надійності ПЗ.

Наведено класифікацію моделей оцінювання надійності ПЗ. В основу значної кількості моделей надійності ПЗ, предметом дослідження яких є кількість помилок у визначеному часовому інтервалі, покладено розподіл Пуассона. Показано, що в моделях цього типу або зовсім не враховується такий показник, як складність програмного проекту, або враховується тільки в часткових випадках (має апріорно встановлені значення, не є параметром моделі), що збільшує похибку оцінювання надійності реальних програмних продуктів (особливо надвеликих програмних систем, де обчислювальна складність алгоритмів оцінювання надійності ПЗ стає неполіноміальною).

Обґрунтовано потребу розроблення нових моделей, методів та засобів для оцінювання та прогнозування надійності ПЗ, які більш адекватно описують процес визначення надійності ПЗ та враховують такий показник, як складність програмного продукту, а також потребу формалізації критерію достатності процесу тестування ПЗ з точки зору його надійності та наявності дефектів. Такий критерій повинен слугувати кількісною мірою для прийняття рішення щодо достатності чи потреби в продовженні процесу тестування. У результаті проведеного аналізу сформульовано основні завдання дисертаційного дослідження.

У **другому розділі** удосконалено моделі надійності ПЗ, де вперше введено показник складності програмного проекту, який є параметром моделі і приймає значення додатного числа. Проведено порівняльний аналіз існуючих та удосконаленої моделей на реальних тестових прикладах та показано переваги моделі з динамічним показником складності проекту над S-подібною моделлю та моделлю Goel–Okumoto. Здійснено формалізацію критерію достатності процесу тестування програмного забезпечення на основі моделі з динамічним показником складності проекту.

Під моделлю надійності ПЗ розуміємо функцію кумулятивної кількості несправностей та функцію інтенсивності виявлення помилок. При побудові моделі оцінювання та прогнозування надійності ПЗ кількість виявлених помилок розподілена за законом Пуассона. Динамічний показник складності проекту є параметром моделі та визначається на основі експериментальних даних і набуває значення додатного числа.

Функція інтенсивності виявлення несправностей ПЗ прийме вигляд:

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

де α – коефіцієнт, який характеризує загальну кількість несправностей, які були виявлені в програмі від початку спостереження, β – коефіцієнт, що характеризує швидкість зміни функції інтенсивності виявлення несправностей, (завжди $\beta > 0$), s – динамічний показник складності проекту.

Для виразу (1) функція кумулятивної кількості помилок ПЗ має вигляд:

$$\mu(t) = \int_0^t \lambda(\tau)d\tau = \alpha\left[-\beta^{s}t^{s}e^{-\beta t} + s\Gamma_{\beta t}(s)\right], \qquad (2)$$

де $\Gamma_z(p) = \int_0^z t^{p-1}e^{-t}dt, \quad (\mathrm{Re}\, p > 0)$ – неповна гамма-функція.

Загальна кількість помилок в ПЗ визначається кумулятивною функцією при $t \to \infty$, таким чином:

$$\mu(\infty) = \alpha s\Gamma(s), \qquad (3)$$

де $\Gamma(s)$ – гамма-функція.

Отже, аналітичний вигляд описаної виразами (1) та (2) моделі дає змогу узагальнити вираз для загальної кількості помилок в (3), яка залежить від величини та складності проекту і визначається параметрами моделі.

Вирази (1) та (2) є моделлю з динамічним показником складності проекту.

Під складністю проекту розуміємо комплексний показник, який пов’язаний з метриками складності коду програмного продукту. На підставі проведених досліджень встановлено орієнтовні інтервали для динамічного показника складності програмного проекту.

Для визначення точкових оцінок параметрів α, β, s моделі застосовано метод максимальної правдоподібності.

Якщо на проміжках $(t_i, t_{i+1}], i = \overline{0,n}$ виявлено m_i помилок, тоді $\sum_{i=1}^{n} m_i = k$. Припустивши, що кількість виявлених помилок розподілена за Пуассоном, та врахувавши інтерпретацію $\mu(t)$ з формули (2), побудовано функцію $L(\alpha,\beta,s)$ максимальної правдоподібності:

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

Функція максимальної правдоподібності L задовольняє такі умови:

- диференційовна при довільних значеннях вибірки;
- досягає максимуму в інтервалі можливих значень.

Значення оцінок величин α, β, s дає розв'язок системи рівнянь:

$$\begin{cases} \dfrac{\partial L(\alpha,\beta,s)}{\partial \alpha} = 0, \\ \dfrac{\partial L(\alpha,\beta,s)}{\partial \beta} = 0, \\ \dfrac{\partial L(\alpha,\beta,s)}{\partial s} = 0. \end{cases} \tag{4}$$

Після виконання потрібних перетворень система диференційних рівнянь (4) набула вигляду:

$$\begin{cases} \alpha = \dfrac{k}{s\Gamma_{\beta t_n}(s) - \beta^s t_n^s e^{-\beta t_n}}; \\ \displaystyle\sum_{i=1}^{n} \frac{m_i\left(t_i^{s+1}e^{-\beta t_i} - t_{i-1}^{s+1}e^{-\beta t_{i-1}}\right)}{s\Phi_{\beta t_i t_{i-1}}(s) + \beta^s\left(t_{i-1}^s e^{-\beta t_{i-1}} - t_i^s e^{-\beta t_i}\right)} - \frac{k t_n^{s+1} e^{-\beta t_n}}{s\Gamma_{\beta t_n}(s) - \beta^s t_n^s e^{-\beta t_n}} = 0; \\ \displaystyle\sum_{i=1}^{n} \frac{m_i}{s\Phi_{\beta t_i t_{i-1}}(s) + \beta^s\left(t_{i-1}^s e^{-\beta t_{i-1}} - t_i^s e^{-\beta t_i}\right)} [\Phi_{\beta t_i t_{i-1}}(s) + s\mathrm{F}_{\beta t_i t_{i-1}}(s) + \\ +\beta^s\left(t_{i-1}^s e^{-\beta t_{i-1}} \ln(\beta t_{i-1}) - t_i^s e^{-\beta t_i}\ln(\beta t_i)\right)] + \\ +\dfrac{k}{s\Gamma_{\beta t_n}(s) - \beta^s t_n^s e^{-\beta t_n}}\left[\beta^s t_n^s e^{-\beta t_n}\ln(\beta t_n) - \Gamma_{\beta t_n}(s) - s\int_0^{\beta t_n} \tau^{s-1}\ln\tau e^{-\tau} d\tau\right] = 0, \end{cases} \tag{5}$$

де $\Phi_{\beta t_i t_{i-1}}(s) = \int_{\beta t_{i-1}}^{\beta t_i} \tau^{s-1} e^{-\tau} d\tau$, $F_{\beta t_i t_{i-1}}(s) = \int_{\beta t_{i-1}}^{\beta t_i} \tau^{s-1}\ln\tau e^{-\tau} d\tau$.

Система трансцендентних рівнянь (5) розв'язана чисельно (модифікованим методом Ньютона), а отримані наближені значення α, β, s використано для аналізу достатності тестування, складності проекту, надійності програмного продукту.

Для верифікації моделі використано експериментальні дані тестування ПЗ, що дало можливість порівняти результати, отримані з використанням розробленої моделі та інших моделей.

Дослідження проведено для моделі (2), моделі Goel–Okumoto та S-подібної моделі, а їх результати наведено в табл.1, де R^2 – квадрат коефіцієнта кореляції між експериментальною та обчисленою кумулятивною кількістю помилок, Δ^2 – середнє квадратичне відхилення.

Таблиця 1.

Значення параметрів опису експериментальних даних дослідженими моделями

Модель	R^2	Δ^2
Модель Goel-Okumoto	0,993	0,12
S-подібна модель	0,995	0,18
Модель з динамічним показником складності проекту	0,999	0,05

Дані табл. 1 показують, що запропонована модель з динамічним показником складності проекту суттєво краще описує реальні експериментальні результати: квадрат коефіцієнта кореляції між експериментальними даними і моделлю з індексом проекту становить 0,999, в той час як для моделі Goel–Okumoto – 0,993, а для S-подібної моделі – 0,995; середнє значення квадрату відхилень між обчисленими та експериментальними значеннями становить 0,05 для моделі (2), що майже на порядок краще ніж двох інших моделей – 0,12 для моделі Goel–Okumoto та 0,18 для S-подібної моделі.

Найвиразніше відмінності між цими трьома моделями проявляються на залежностях функції інтенсивності виявлення помилок від часу, які наведено на рис. 1 (крива 1 – модель з динамічним показником складності проекту, крива 2 – модель Goel–Okumoto, крива 3 – S-подібна модель).

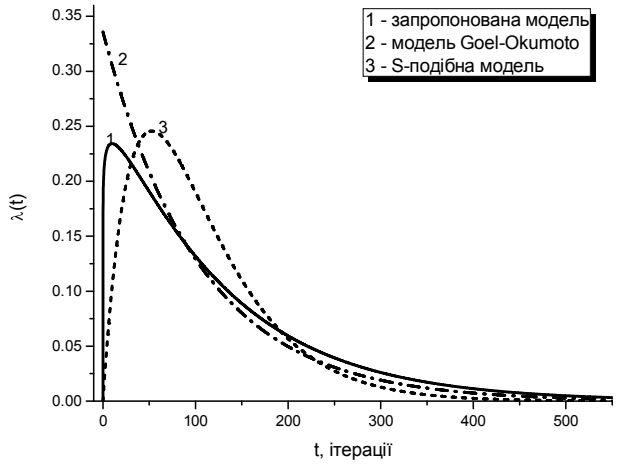

Рис.1. Залежність $\lambda(t)$ від кількості ітерацій для різних моделей

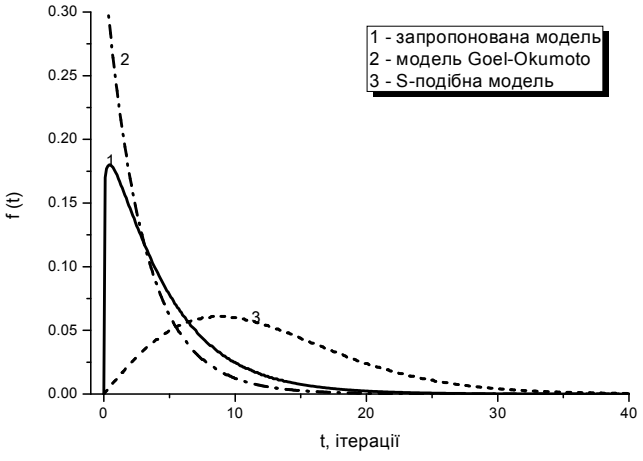

Рис.2. Залежність частоти помилок *f(t)* від кількості ітерацій для різних моделей

Як показано на рис. 1, модель Goel–Okumoto, як і S-подібна модель показують завищені, порівняно з даними експерименту, значення інтенсивності виявлення помилок при малих значеннях часу, і навпаки – занижені значення

при великих t, вказуючи на більш раннє завершення процесу виявлення помилок, ніж спостерігається в реальних експериментальних даних. Крім того, модель Goel–Okumoto неточно описує початковий етап тестування (рис.1). Натомість S-подібна модель описує середину процесу тестування з найбільшим відхиленням (в нашому випадку приблизно від 50-ї до 175-ї ітерації).

Графік залежності частоти помилок від часу для трьох розглянутих моделей надійності наведено на рис. 2.

Як і у випадку функції інтенсивності виявлення помилок, залежність частоти помилок (рис. 2) дає підстави зробити аналогічні висновки. Модель Goel–Okumoto (крива 2, рис. 2) показує, що частота помилок є аномально високою відносно експериментальних даних в початкові моменти часу і аномально швидко спадає з часом, що не відповідає практичним результатам тестування ПЗ. Натомість S-подібна модель (крива 3, рис. 2) показує занижену максимальну частоту помилок при більш повільному її зниканні з часом.

З метою формалізації критерію достатності процесу тестування ПЗ побудовано графічну залежність поведінки динамічного показника складності проекту s, яка виявляє чітку особливість, що покладена в основу критерію достатності процесу тестування (рис. 3). Ця особливість полягає в тому, що при $t \geq 550$ (що відповідає переходу до пуассонового розподілу кількості виявлених помилок) залежність $s(t)$ стає гладкою, а значення s наближається до сталої величини, на відміну від такої залежності при $t < 550$, яка проявляє різко осцилюючий характер.

Така поведінка залежності $s(t)$ пояснюється тим, що динамічний показник складності проекту (параметр s моделі) є основним параметром, який визначає форму і функцію розподілу, а відповідно і густину розподілу імовірностей випадкової величини, яка у нашому випадку визначається кількістю виявлених помилок.

Отже на пізніх етапах тестування ПЗ, коли взаємозалежні помилки виявлені та усунені, а кількість виявлених помилок відповідає пуассоновому розподілу, параметр s вже практично не змінюється, а змінюються в основному кількісні характеристики (параметри α та β), що дає можливість формалізувати критерій достатності процесу тестування ПЗ таким чином:

$$\frac{ds(t)}{dt} \xrightarrow[t \to \infty]{} 0 \qquad (6)$$

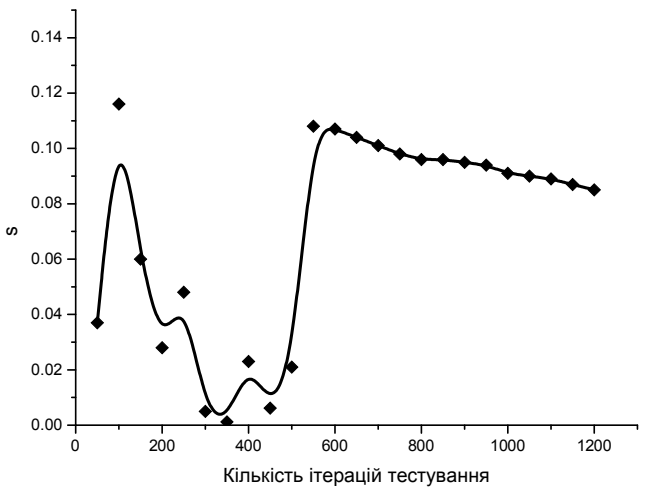

Рис.3. Залежність параметру s побудованої моделі від кількості ітерацій.

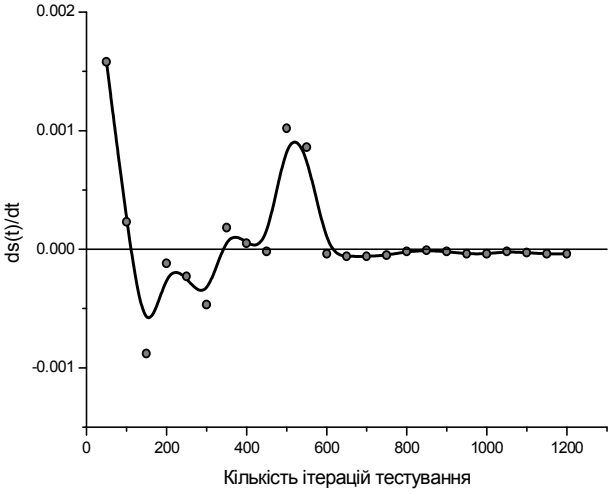

Рис.4. Залежність критерію достатності процесу тестування ПЗ від кількості ітерацій.

Залежність критерію (6), отриману шляхом чисельного диференціювання даних з рис. 3, наведено на рис. 4. Така залежність ілюструє критерій достатності процесу тестування ПЗ (6) і показує різку зміну форми кривої при переході експериментальних даних до пуассонового розподілу. Таким чином, є змога з використанням критерію достатності процесу тестування визначити загальну кількість помилок в програмному проекті за допомогою рівняння (3) і, порівнявши її з кількістю вже виявлених та виправлених помилок, прийняти обґрунтоване рішення про розподіл ресурсів проекту.

У **третьому розділі** проведено дослідження основних припущень та обмежень моделі надійності ПЗ з динамічним показником складності проекту, які доповнюють математичний формалізм моделі.

Проведено обґрунтування та сформульовані рекомендації щодо застосування моделі в умовах реального циклу виробництва програмного забезпечення. Розроблено метод оцінювання та прогнозування надійності ПЗ на основі кількості виявлених помилок та процедуру підтримки прийняття рішень на основних стадіях життєвого циклу ПЗ, базуючись на новій моделі. Застосування методу ілюструється прикладом з використанням експериментальних даних тестування промислового програмного продукту. Ілюстрація послідовності дій цього методу наведена на рис. 5 і складається з 7 кроків.

Крок 1. Отримання та підготовка даних про помилки ПЗ. Перед початком застосування моделі слід забезпечити дотримання вимог та спрощень моделі і ретельно проаналізувати вхідні дані. Необхідно представити їх у вигляді кількості помилок на вибраних часових інтервалах.

Крок 2. Визначення початкових наближень параметрів моделі. Оскільки параметри моделі отримуються як розв'язки системи рівнянь (5) ітераційним

методом, треба встановити початкові наближення для параметрів моделі α, β, s. Необхідно зазначити, що їх можливо встановити тільки після проведення тестування хоча б на одному часовому інтервалі

Крок 3. Обчислення параметрів моделі. Чисельно розв’язати систему рівнянь (5). Для розв'язання цієї системи використовується модифікований метод Ньютона з початковими наближеннями, отриманими на кроці 2. Маючи розв'язок системи рівнянь, обчислені точкові оцінки підставляють у модель і отримують числовий опис моделі на основі наявних даних про поведінку помилок.

Крок 4. Перевірка якості опису експериментальних даних моделлю. Для цього необхідно:

- провести тест Колмогорова–Смірнова;
- обчислити значення квадрату коефіцієнта кореляції Пірсона між експериментальною вибіркою та теоретичною вибіркою;
- знайти середнє квадратичне відхилення.

Якщо згадані тести проведено успішно – експериментальні дані з достатньою точністю описуються запропонованою моделлю і можна продовжувати проводити оцінювання та прогнозування надійності ПЗ. У іншому випадку слід змінити початкові значення (крок 2) і отримати нові точкові оцінки моделі (крок 3), або ж характер процесу виявлення помилок.

Крок 5. Перевірка критерію достатності процесу тестування. Для перевірки цього критерію слід мати набір значень параметру s в різні моменти часу (не менше трьох точок). Будується залежність $s(t)$, де в ролі абсцис будуть значення $t_1, t_2, ..., T$, а відповідними ординатами – точкові оцінки параметру s, отримані з розв'язку системи рівнянь максимальної правдоподібності до експериментальних даних з відповідного часового інтервалу. Визначена залежність чисельно диференціюється. Критерієм достатності процесу тестування буде зменшення похідної до нуля.

Крок 6. Обчислення кількісних характеристик надійності ПЗ. На цьому етапі обчислюють різні кількісні характеристики для оцінювання надійності програмної системи, такі як прогнозована загальна кількість помилок в системі, кількість залишкових помилок в системі, середній час до виявлення наступної помилки, імовірність безвідмовної роботи тощо.

Крок 7. Процес прийняття рішень. На цьому кроці приймаються такі рішення на основі інформації, отриманої на попередній кроках описаного методу, як: рішення стосовно випуску фінальної версії чи продовження процесу розробки, перерозподілу ресурсів проекту.

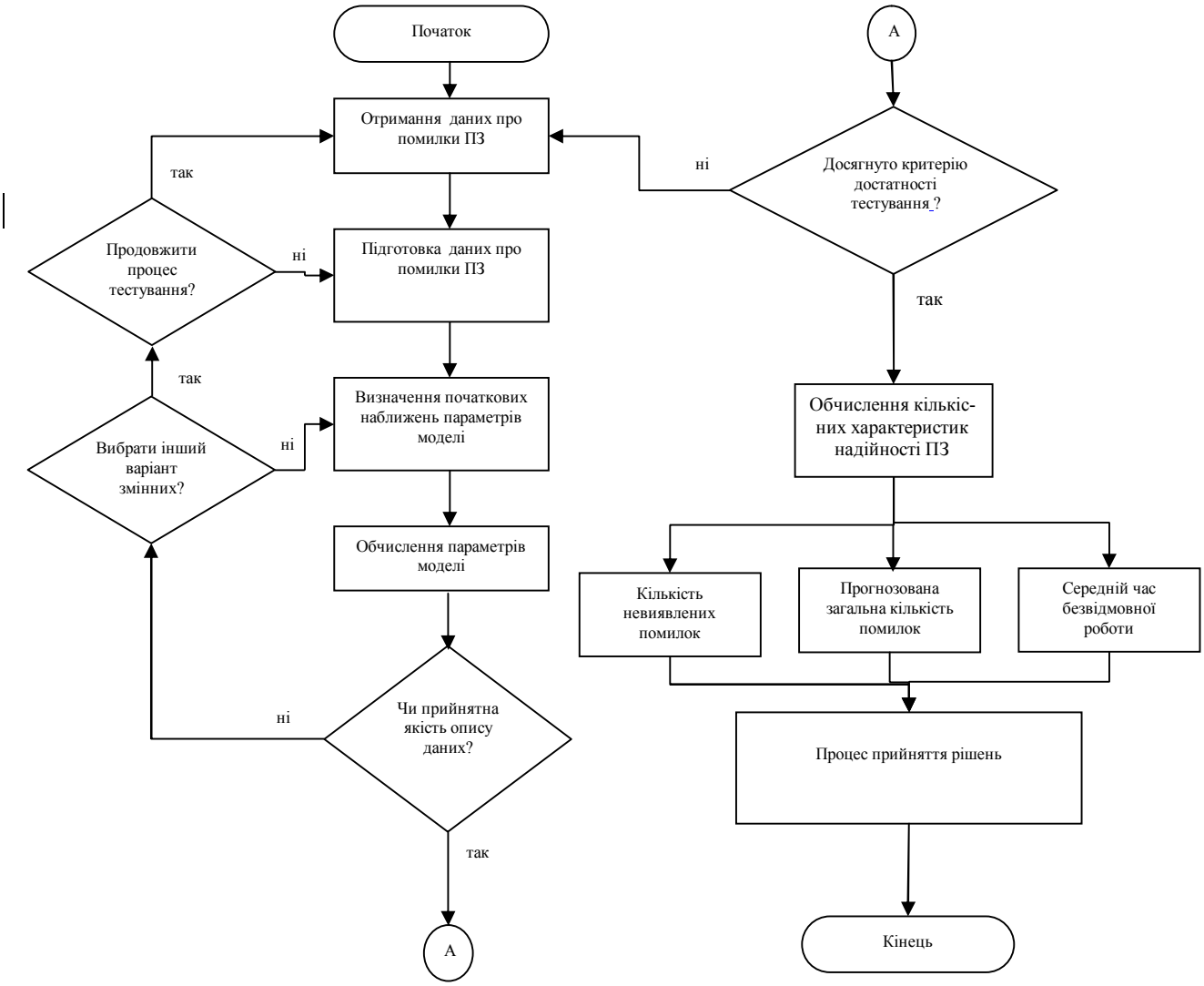

Рис.5. Блок-схема методу оцінювання та прогнозування надійності ПЗ

Перевагою цієї моделі над моделями на основі внесення помилок є відсутність необхідності внесення помилок в код програми; перевагою над моделями на основі часу між помилками є її толерантність до детермінованого процесу тестування, який найчастіше використовується при розробці промислових програмних продуктів, а перевагою порівняно з моделями на основі області вхідних даних є менш строга вимога до передбачення експлуатаційного виконання програми, яка може бути замінена вимогою до тестування усіх модулів програми замість усіх логічних шляхів виконання програми.

Результати експериментів показують, що ця модель завдяки її перевагам над основними відомими моделями такого класу є найбільш придатною для використання на етапі інтеграційного тестування, при цьому слід звернути особливу увагу на використання взаємно незалежних наборів тестових даних та однорідного охоплення процесом тестування принаймні усіх модулів та компонентів програмної системи.

В **четвертому розділі** спроектовано та розроблено програмну систему для оцінювання та прогнозування надійності ПЗ на основі розробленого математичного забезпечення, яка дає змогу автоматизувати опрацювання результатів процесу тестування ПЗ та практичні рекомендації щодо його достатності чи необхідності продовження. Розроблена система має двошарову архітектуру клієнт-сервер (рис.6).

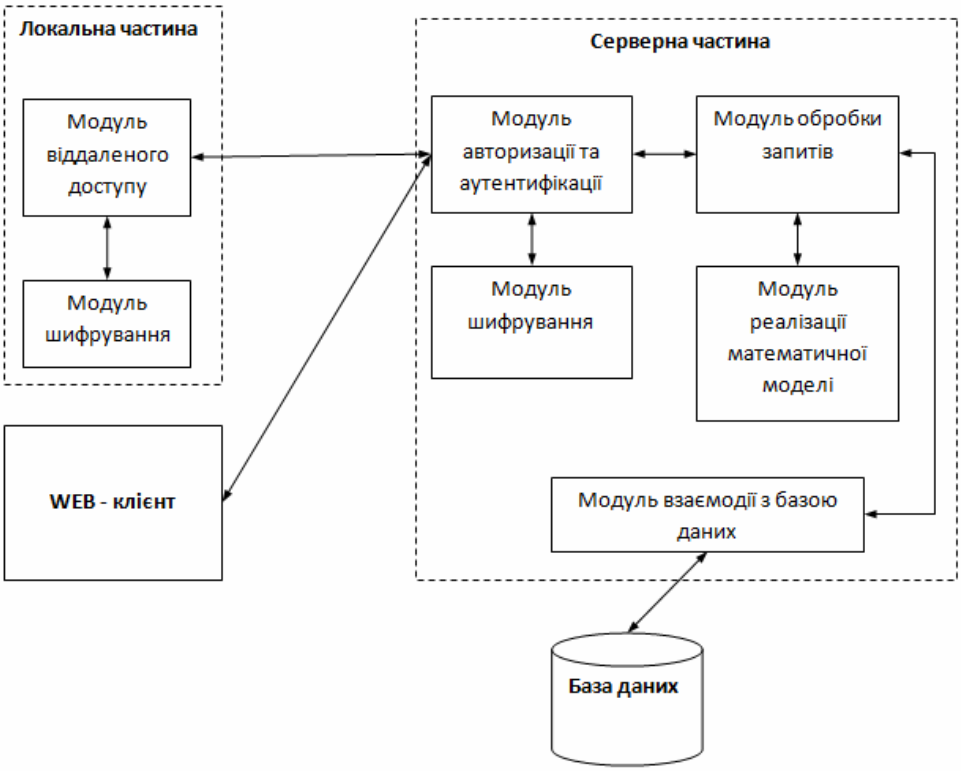

Рис. 6. Архітектура розробленої системи

При використанні локальної частини програмної системи користувач виконує усі обчислення безпосередньо на своєму локальному комп'ютері. Цей підхід актуальний при використанні програмної системи одним користувачем або при неможливості виділення окремого ресурсу під клієнт-серверну версію програмної системи. Локальна частина також може використовуватись в якості клієнта серверної версії. В локальній версії передбачена можливість взаємодії з віддаленим сервером для централізованого збереження даних про досліджуваний програмний продукт та отримання результатів тестування різним типам користувачів (дослідник, розробник, замовник).

Клієнт-серверна частина програмної системи призначена для централізованого збереження даних про досліджуваний програмний продукт та надання користувачам потужних обчислюваних ресурсів серверних станцій. Тут передбачено використання модуля для взаємодії з веб-браузерами, що робить дану програмну систему платформонезалежною та уможливлює взаємодію з користувачем із будь-якого комп'ютера, на якому встановлено веб-браузер.

Створена програмна система являє собою WindowsFormsApplication і дає користувачу можливість взаємодіяти з системою в звичному візуальному інтерфейсі. Для одночасної роботи з декількома електронними документами систему реалізовано у вигляді Multi Document Interface. Нижче (рис. 7) відображено вигляд вікна візуалізації опрацювання експериментальних даних.

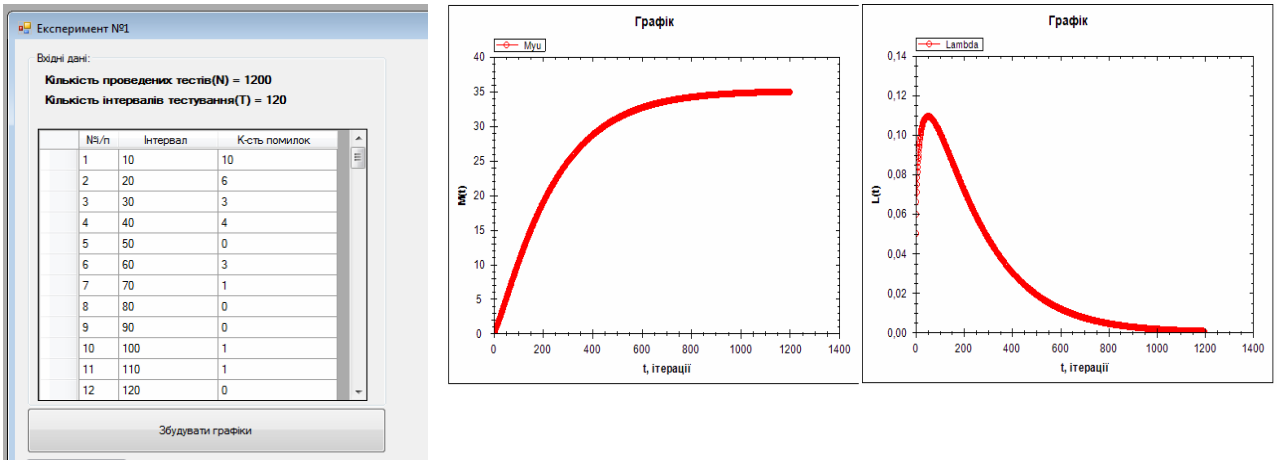

Рис. 7. Фрагменти вікон візуалізації опрацювання експериментальних даних

Проведено дослідження безпеки даних розробленої програмної системи, що зумовлено її клієнт-серверною архітектурою, та обгрунтовано вибір алгоритму захисту даних. Для захисту файлів проектів локальної версії використовується алгоритм симетричного шифрування RC5. Для обміну даними з віддаленим сервером використовується алгоритм шифрування з відкритим ключем RSA.

ОСНОВНІ РЕЗУЛЬТАТИ ТА ВИСНОВКИ

У дисертації розв’язано актуальну наукову задачу створення засобів для прогнозування надійності програмного забезпечення з використанням показника складності програмного проекту.

Основні наукові та практичні результати дисертаційної роботи:

1. Удосконалено математичну модель надійності програмного забезпечення на основі пуассонового розподілу, яка за рахунок уведення динамічного показника складності програмного проекту дає можливість покращити коефіцієнт кореляції між експериментальними даними та результатами моделювання на 0,8–1,2 %, а середньоквадратичне відхилення на 25% порівняно з найбільш поширеними моделями цього класу – S-подібною моделлю та моделлю Goel–Okumoto.

2. Формалізовано критерій достатності процесу тестування ПЗ, який базується на поведінці залежності динамічного показника складності проекту від кількості циклів тестування програмного продукту. Використання цього критерію дає змогу зменшити похибку прогнозування загальної кількості

помилок в ПЗ з 25–30% (критерій достатності не досягнуто) до 2–3% (після досягнення критерію достатності).

3. У результаті аналізу основних припущень та обмежень моделі надійності ПЗ з динамічним показником складності проекту показано шляхи забезпечення коректності використання моделі у випадку, коли деякі з припущень не справджуються в процесі розробки програмного продукту. Для цього слід використовувати взаємно незалежні набори тестових даних та охопити процесом тестування усі модулі та компоненти програмної системи.

4. Розроблена модель надійності ПЗ придатна до використання на підприємстві для оцінювання загальної кількості помилок в ПЗ, прогнозування частоти їх появи та обчислення кількісних показників надійності ПЗ. Динамічний показник складності проекту та значення критерію достатності процесу тестування є підставою для практичних рекомендацій керівникам проекту стосовно розподілу виробничих ресурсів між етапами життєвого циклу проекту.

5. Розроблено метод оцінювання та прогнозування надійності програмного забезпечення з урахуванням значення критерію достатності процесу тестування, який дає змогу зменшити на 10% похибку при визначенні кількісних характеристик надійності ПЗ, зокрема кількості залишкових несправностей.

6. Розроблено програмне забезпечення для оцінювання та прогнозування надійності ПЗ на основі отриманих в роботі теоретичних результатів, яке дає змогу автоматизувати опрацювання експериментальних даних і оцінювати достатність процесу тестування програмних проектів, що забезпечує практичні рекомендації керівникам проекту стосовно розподілу виробничих ресурсів між етапами життєвого циклу розробки ПЗ.

7. Здійснено верифікацію розробленого математичного та програмного забезпечення шляхом впровадження у виробництво розроблених засобів прогнозування надійності програмного проекту із врахуванням показника його складності, які дали можливість зменшити загальні економічні затрати на розробку програмного проекту на 5%, що підтверджується відповідними актами про впровадження.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЙНОЇ РОБОТИ

1. Критерій достатності процесу тестування програмного забезпечення / В. С. Яковина, М. М. Сенів, Я. М. Чабанюк, Д. В. Федасюк, У. Т. Хімка // Комп'ютерні науки та інформаційні технології. – 2010. – С. 346–358. – (Вісник / Нац. ун-т "Львів. політехніка" ; № 672).

2. Аналіз використання моделі надійності програмного забезпечення з динамічним показником складності проекту протягом життєвого циклу / Сенів М. М., Федасюк Д. В., Чабанюк Я. М., Яковина В. С. // Комп'ютерні технології друкарства : зб. наук. пр. / Укр. акад. друкарства. – Львів, 2010. – №.24 – С. 111–126.
3. Сенів М. М. Розробка методу оцінювання та прогнозування надійності програмного забезпечення на основі моделі з врахуванням критерію достатності процесу тестування / Сенів М. М., Федасюк Д. В., Яковина В. С. // Комп'ютерні технології друкарства : зб. наук. пр. / Укр. акад. друкарства. – Львів, 2011. – № 25 – С. 72–83.
4. Програмний модуль прогнозування надійності системи теплового проектування / Д. Федасюк, М. Сенів, П. Сердюк, Н. Мамроха // Комп'ютерні науки та інформаційні технології. – 2008. – С. 161–165. – (Вісник / Нац. ун-т "Львів. політехніка ; № 629).
5. Програмна система оцінювання та прогнозування надійності програмного забезпечення / Сенів М. М., Федасюк Д. В., Парфенюк Ю. І., Яковина В. С., Чабанюк Я. М. // Відбір і обробка інформації : міжвід. зб. наук. пр. / НАН України, Фіз.-мех. ін-т ім. Г.В. Карпенка. – Львів, 2010. – Вип. 33(109). – С. 123–129.
6. Дослідження характеристик криптостійкості алгоритму симетричного шифрування DES / Яковина В. С., Федасюк Д. В., Салій С. І., Сенів М. М. // Комп'ютерні системи проектування. Теорія і практика. – Львів : Вид-во Нац. ун-ту "Львів. політехніка", 2008. – С. 55–62. – (Вісник / Нац. ун-т "Львівська політехніка" ; № 626).
7. Дослідження основних характеристик алгоритму симетричного шифрування RC5 для побудови модуля захисту розподіленої системи теплового проектування / Яковина В. С., Одуха О. В., Сенів М. М., Білас О. Є. // Комп'ютерні науки та інформаційні технології. – Львів : Вид-во Нац. ун-ту "Львів. політехніка", 2008. – С. 143–150. – (Вісник / Нац. ун-т "Львівська політехніка" ; № 616).
8. Порівняння швидкодії програмної реалізації алгоритмів симетричного (DES) та асиметричного (RSA) шифрування / Яковина В., Федасюк Д., Сенів М., Білас О. // Комп'ютерні науки та інформаційні технології. – Львів : Вид-во Нац. ун-ту "Львів. політехніка", 2007 – С. 181–185. – (Вісник / Нац. ун-т "Львів. політехніка" ; № 598).
9. Побудова і дослідження моделі надійності програмного забезпечення з індексом величини проекту / Я. М. Чабанюк, В. С. Яковина, Д. В. Федасюк , М. М. Сенів, У. Т. Хімка // Інженерія програмного забезпечення. – 2010. – № 1. – С. 24–29.
10. Architecture of software system for the software reliability evaluation and prediction / Maksym Seniv, Dmytro Fedasyuk, Yuriy Parfenyuk, Vitaliy Yakovyna, Yaroslav Chabanyuk // Комп'ютерні науки та інформаційні технології : матеріали П'ятої міжнар. наук.-техн. конф. CSIT-2010, 14–16 жовт. 2010 р., Львів. – Львів, 2010. – С. 164–167.

11. The estimation and prediction model of the software reliability with the project size index / D. Fedasyuk, M. Seniv, Y. Chabanyuk, U. Khimka // Proc. of the Xth Intern. conf. “Modern problems of radio engineering, telecommunications, and computer science”, Lviv, Slavske, Febr. 23–27, 2010. – Lviv, 2010. – P. 209–210.
12. Information model of data representation for software reliability estimation framework / Fedasyuk D. V., Seniv M. M., Yakovyna V. S., Mamrokha N. M. // Proc. of the Xth Intern. conf. “The experience of designing and application of CAD Systems in Microelectronics” CADSM 2009, Lviv, Polyana, 24–28 Febr. 2009. – Lviv, 2009. – P. 287–291.
13. Побудова моделі аспекту аутентифікації та авторизації для підсистеми захисту програмних систем / Дмитро Федасюк, Віталій Яковина, Максим Сенів, Назар Мамроха // Комп'ютерні науки та інформаційні технології : матеріали Четвертої міжнар. наук.-техн. конф. CSIT-2009, 15–17 жовт. 2009 р., Львів. – Львів, 2009. – С. 198–202.
14. Яковина В. Аспектна декомпозиція компонентів захисту розподіленої системи теплового проектування / Віталій Яковина, Назарій Мамроха, Максим Сенів // Зб. матеріалів Шостої міжнар. конф. "Інтернет – Освіта – Наука – 2008" ІОН-2008, 7–11 жовт. 2008 р., Вінниця. – Вінниця : УНІВЕРСУМ-Вінниця, 2008. – Т. 2. – С. 407–410.
15. Fedasyuk D. The reliability analysis of thermal design software system / Dmyrtro Fedasyuk, Maksym Seniv, Pavlo Serdyuk // Proc. of the IVth Intern. conf. of young scientists MEMSTECH’2008, 21–24 May, Polyana, Ukraine. – Lviv, 2008. – P. 122.
16. Yakovyna V. D. Software realization and performance testing of DES cryptographic algorithm on the .NET platform / V. Yakovyna, D. Fedasyuk, M. Seniv // Proc. of the IXth Intern. conf. “The experience of designing and application of CAD Systems in Microelectronics” CADSM 2007, Lviv, Polyana, Ukraine, 20–24 Febr. 2007. – Lviv, 2007. – P. 386–388.
17. The performance testing of RSA algorithm software realization / V. Yakovyna, D. Fedasyuk, M. Seniv, O. Bilas // Proc. of the IXth Intern. conf. “The experience of designing and application of CAD systems in microelectronics” CADSM 2007, Lviv, Polyana, 20–24 Febr. 2007. – Lviv, 2007. – P. 390–392.
18. Федасюк Д. В. Розробка програмного модуля оцінки надійності програмного забезпечення на основі моделі Нельсона / Д. В. Федасюк, М. М. Сенів, Н. М. Мамроха // Комп’ютерні науки та інженерія : матеріали 2-ої Міжнар. конф. молодих науковців CSE-2007, 4–6 жовт. 2007, Львів. – Львів : Вид-во Нац. ун-ту «Львів. політехніка», 2007. – С. 17–19.

АНОТАЦІЇ

Сенів М. М. Засоби прогнозування надійності програмного проекту із врахуванням показника його складності. – Рукопис.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 01.05.03 – математичне та програмне забезпечення обчислювальних машин і систем. – Національний університет "Львівська політехніка", Львів, 2011.

Дисертація присвячена розробленню засобів для прогнозування надійності програмного забезпечення з використанням показника складності програмного проекту. Удосконалено математичну модель надійності програмного забезпечення на основі пуассонового розподілу, використання якої дає можливість покращити коефіцієнт кореляції між експериментальними даними та результатами моделювання на 0,8–1,2 %, а середньоквадратичне відхилення на 25% порівняно з найбільш поширеними моделями цього класу. Описано та формалізовано критерій достатності процесу тестування ПЗ, використання якого дає змогу зменшити похибку прогнозування загальної кількості помилок в ПЗ з 25–30% (критерій достатності не досягнуто) до 2–3 % (після досягнення критерію достатності). Розроблено метод оцінювання та прогнозування надійності програмного забезпечення, який дає змогу зменшити на 10% похибку при визначенні кількісних характеристик надійності ПЗ.

Розроблена програмна система, яка реалізує запропоноване математичне забезпечення та дає змогу автоматизувати обробку експериментальних даних і оцінювати достатність процесу тестування програмних продуктів, що дає можливість надати практичні рекомендації керівникам проекту стосовно розподілу виробничих ресурсів між етапами життєвого циклу розробки ПЗ.

*Ключові слова***:** надійність програмного забезпечення, моделі оцінювання надійності ПЗ, показник складності проекту, критерій достатності тестування, достатність процесу тестування, життєвий цикл ПЗ.

Сенив М.М. Средства прогнозирования надежности программного проекта с учетом показателя его сложности. – Рукопись.

Диссертация на соискание ученой степени кандидата технических наук по специальности 01.05.03 - математическое и программное обеспечение вычислительных машин и систем. - Национальный университет "Львовская политехника", Львов, 2011.

Диссертация посвящена разработке средств для прогнозирования надежности программного обеспечения с использованием показателя сложности программного проекта. Усовершенствована математическая модель надежности программного обеспечения на основе пуассонова распределения, использование которой дает возможность улучшить коэффициент корреляции

между экспериментальными данными и результатами моделирования на 0,8-1,2%, а среднеквадратичное отклонение на 25% по сравнению с наиболее распространенными моделями этого класса. Описан и формализован критерий достаточности процесса тестирования ПО, использование которого позволяет уменьшить погрешность прогнозирования общего количества ошибок в ПО с 25-30% (критерий достаточности не достигнут) до 2-3% (по достижении критерия достаточности). Разработан метод оценки и прогнозирования надежности программного обеспечения позволяющий уменьшить на 10% погрешность при определении количественных характеристик надежности ПО.

Разработана программная система, реализующая предложенное математическое обеспечение и дает возможность автоматизировать обработку экспериментальных данных и оценивать достаточность процесса тестирования программных продуктов, что дает возможность предоставить практические рекомендации руководителям проекта по распределению производственных ресурсов между этапами жизненного цикла разработки ПО.

*Ключевые слова***:** надежность программного обеспечения, модели оценки надежности ПО, показатель сложности проекта, критерий достаточности тестирования, достаточность процесса тестирования, жизненный цикл ПО.

Seniv M.M. Software project reliability prediction tools subject to its complexity index. - Manuscript.

Thesis for Ph.D. degree in the speciality 01.05.03 - Mathematical and software support of computer machines and systems. - Lviv Polytechnic National University , Lviv, 2011.

The thesis is dedicated to the development of tools for software reliability prediction using the software project complexity index. The classification of software reliability estimation models has been introduced. The Poisson distribution is the underlying basis of a large number of software reliability models, which have errors in a defined time interval as the research subject. It has been shown that in models of this type the software project complexity index is either not taken into account, or considered only in partial cases (has a priori value, is not a model parameter).This increases the inaccuracy of reliability estimation for the real software products (especially very large-scale software systems where the computational complexity of algorithms for software reliability estimation becomes nonpolynomial).

The necessity in the development of new models, methods and tools for the software reliability estimation and prediction, which more adequately describe the process of software reliability evaluation and take into account the software complexity index and also the need to formalize the criterion of the software testing process adequacy in terms of its reliability and defects occurrence. This criterion should serve as a quantitative measure for making the decision on the adequacy or the need for the testing process continuation.

The mathematical model of software reliability based on Poisson distribution has been improved. Its application allows to improve the correlation coefficient between experimental data and modeling results to 0,8-1,2% and standard deviation to 25% compared with the most common models of this class. A comparative analysis of existing and improved models to real test cases has been conducted and the advantages of models with dynamic index of project complexity over S-shaped model and the Goel-Okumoto model have been shown.

In order to formalize the criterion of software testing process adequacy the graphical representation of behavior of the project complexity dynamic index s has been built, which reveal a clear feature that can be underlying for the testing process adequacy criterion. This feature lies in the following, that during the transition to Poisson distribution of revealed mistakes number the dependence $s(t)$ becomes differentiable and the value s approaches constant value, contrary to the dependence on the transition point, which shows sharply oscillating character. The usage of the software testing process adequacy criterion allows to reduce the prediction inaccuracy of the total number of errors in the software from 25-30% (adequacy criterion is not achieved) to 2-3% (after achieving the adequacy criterion).

It has been proved and formulated guidelines for a model usage in a real cycle of software production. The method of software reliability estimation and prediction based on the number of revealed mistakes and the support procedure of decision making at key stages of the software life cycle based on the new model have been developed. The application of the method is illustrated by example using experimental data of industrial software product testing, which consists of 7 steps. The developed method of the software reliability estimation and prediction allows reducing the inaccuracy in determining the software reliability quantitative characteristics by 10%.

The software system that implements the proposed mathematical support has been developed, and it allows to automatize the experimental data processing and to assess the adequacy of the software product testing process, that gives the opportunity to provide project managers with practical recommendations concerning the productive resources distribution between stages of the software development life cycle. The developed system has two-layer client-server architecture.

The data security investigation of the developed software system, due to its client-server architecture, has been conducted and the choice of data protection algorithm has been proved. The symmetric encryption algorithm RC5 has been used to protect the project files of local version. To communicate with a remote server the RSA public key encryption algorithm has been used.

Keywords: software reliability, software reliability estimation model, project complexity index, testing adequacy criterion, testing process adequacy, software life cycle.