УЗАГАЛЬНЕНА МОДЕЛЬ НЕ ГОМОГЕННОГО ПУАСОНІВСЬКОГО ПРОЦЕСУ ДЛЯ ОЦІНЮВАННЯ НАДІЙНОСТІ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

Тимошенко Ю.О., Дідковська М.В.

Навчально-науковий комлекс „Інститут прикладного системного аналізу" Національного технічного університету „КПІ", 03056, м. Київ, пр. Перемоги, 37, корп. 35, тел/факс.236-39-87, e-mail: YTim@unibox.com.ua
Дана стаття присвячена питанням оцінювання надійності програмного забезпечення, дослідженню моделей, що базуються на не гомогенному пуасонівському процесі та розробці узагальненої моделі.

This article is devoted to the reliability estimation's theory and non homogeneous Poisson process models investigation. The generalized finite non Homogeneous Poisson model is proposed.

Вступ

Обчислювальні системи, невід'ємною частиною яких є програмне забезпечення (ПЗ), відіграють важливу роль в усіх галузях нашого життя.

Розроблення ПЗ – це складний, дорогий та довготривалий процес. Його можна вважати завершеним лише при досягненні ПЗ специфікованого рівня якості. Вочевидь, що досягнення цього рівня вчасно є критичним при розробці програмних продуктів, які застосовуються в банківських системах, авіоніці, управлінні ядерними реакторами, медицині тощо. Якість ПЗ має магато атрибутів, одним з найважливіших з яких є надійність.

Таким чином, задача оцінювання надійності ПЗ стає найактуальнішою на сьогоднішній день [1].

Надійність ПЗ може бути оцінена за допомогою статичних та динамічних моделей або підтвердження шляхом статистичного тестування.

В рамках даної роботи автори розглядають лише один клас динамічних моделей оцінювання надійності, а саме моделей, що базуються на негомогенному пуасонівському процесі, аналізують їх практичне застосування та властивості та пропонують узагальнену модель.

Математичні основи надійності.

Існує багато конкуруючих поглядів на те, що називається надійністю ПЗ. Один з таких поглядів полягає в тому, що програма або коректна або некоректна [2], тоді надійність ПЗ – бінарна за своєю природою: ідеальна програма має надійність рівну одиниці, а неідеальна – рівну нулеві. Інший підхід базується на тестуванні ПЗ, в цьому випадку процентне відношення вдалих тестів використовується для вимірювання надійності, тобто надійність ПЗ визначена як відносна частота вдалих виконань програми[2]. В [3,4] надійність визначена як ймовірність безвідмовного функціонування ПЗ в специфікований час чи в специфікованому середовищі. Надалі використовується саме це визначення надійності ПЗ.

Розглянемо математичні основи надійності [3].

Ймовірність того, що в інтервалі часу від t до t+Δt виникне несправність дорівнює:

$$P(t \le T \le t+\Delta t) = f(t)\Delta t = F(t+\Delta t) - F(t),$$

де F(t) - функція розподілу несправностей, а f(t) – функція щільності. Тоді функція надійності матиме вигляд:

$$R(t) = P(T > t) = 1 - F(t) = \int_t^{\infty} f(x)dx .$$

Інтенсивність несправностей – ймовірність виникнення несправності в інтервалі часу від t до t+Δt, на одиницю часу

$$\frac{P(t \le T \le t+\Delta t \,|\, T > t)}{\Delta t} = \frac{P(t \le T \le t+\Delta t)}{\Delta t P(T>t)}$$

$$= \frac{F(t+\Delta t) - F(t)}{\Delta t R(t)}.$$

Розглянемо границю функції інтенсивності відмов за умови, що часовий інтервал прямує до нуля:

$$z(t) = \lim_{\Delta t \to 0} \frac{F(t+\Delta t) - F(t)}{\Delta t R(t)} = \frac{f(t)}{R(t)} .$$

Функцію надійності можна виразити наступним чином:

$$R(t) = \exp(-\int_0^t z(x)dx)$$

Нехай M(t) - випадковий процес, який характеризує кумулятивну кількість відмов на момент часу t.

Позначимо μ(t) як його математичне сподівання:

$\mu(t) = E(M(t))$, тоді функція інтенсивності несправностей є похідною від μ(t): $\lambda(t) = \dfrac{d\mu(t)}{dt}$

Під час розробки та тестування ПЗ стає необхідним оцінювати його надійність на практиці.

Моделі оцінювання надійності ПЗ

Для прогнозування надійності ПЗ використовують різноманітні аналітичні методи [3]. Існує багато методів оцінювання надійності програмного забезпечення, вони можуть бути розділені на два основних класи – за областю визначення даних (статичні) та часу (динамічні), які отримали назву “моделі зростання надійності”.

Динамічні моделі аналізують відмову як процес в той час, як статичні намагаються аналізувати несправності як вміст програмного продукту. Використання статичних моделей для отримання точних оцінок надійності програмного забезпечення дуже дороге, в той час як моделі зростання несправностей – економічно більш вигідні. Динамічні моделі також найкраще задовольняють вимогам перевірки проектів зі щільним графіком написанням або проектів з простою структурою програми. Проте у випадку програм зі складною структурою, таких як системне ПЗ та комунікаційні програми, використання цих моделей призводить до недооцінки кількості наявних несправностей. Моделі зростання несправностей оцінюють та прогнозують загальну надійність, в той час коли дані отримані за допомогою статичних моделей можуть бути використані для ґрунтовного тестування програмних продуктів.

В рамках даної статті ми будемо розглядати лише динамічні моделі, більш того лише моделі, що базуються на негомогенному пуасонівському процесі.

Динамічні моделі

Мета моделей даного класу полягає в дослідженні розвитку надійності ПЗ у часі. Ці динамічні моделі отримали назву моделей зростання надійності [3] та були представлені їх авторами через математичне сподівання функції кумулятивної кількості відмов на момент часу t - μ(t) та функцію інтенсивності несправностей λ(t).

В даній роботі вважається за доцільне додатково розглянути для кожної з моделей також її функцію надійності R(t) з метою - полегшити дослідження надійності ПЗ у часі та визначити її граничні значення. Розглянемо моделі зростання надійності, що базуються на не гомогенному пуасонівському процесі, більш детально.

Модель не гомогенного пуасонівського процесу – модель Гоеля-Окумото

Ця модель була вперше запропонована у 1979 Гоелем та Окумото [5]. В цій моделі вважається, що кількість відмов за одиницю часу - незалежні пуасонівські випадкові величини.

Припущення моделі:

Враховуючи базові, модель має наступні припущення:

1. Кумулятивна кількість відмов на момент t, M(t) відповідає пуасонівському процесу з функцією математичного сподівання μ(t). Мат. сподівання таке, що очікувана кількість несправностей, які буде виявлено за проміжок часу t+Δt пропорційна кількості несправностей, що ще залишилися у системі на момент t. Також вважаємо, що μ(t) – обмежена, не спадаюча функція часу: $\lim_{t\to\infty} \mu(t) = N < \infty$. Отже, ця модель належить до категорії скінчених функцій.
2. Кількості несправностей ($f_1, f_2, \ldots, f_n$), виявлених на відповідних часових інтервалах [(t_0=0,t_1), (t_1,t_2),…, (t_{n-1},t_n)] незалежні для будь-якої скінченої кількості часових проміжків $t_1<t_2<\ldots<t_n$.

Для моделі негомогенного пуасонівського процесу використовують такі набори даних:

f_i - кількість несправностей на кожному з досліджуваних інтервалів;

t_i – моменти закінчення досліджуваних часових інтервалів.

Формальний вигляд моделі

Враховуючі припущення моделі та її опис, можна записати:

$$\mu(t) = N(1-\exp(-bt)),$$

константи *N>0, b>0. N* - очікувана загальна кількість несправностей, які будуть виявлені.

Тоді функція інтенсивності відмов матиме вигляд:

$$\lambda(t) = Nb\exp(-bt)$$

Зауважимо, що ця функція строго спадає при t>0.

Виходячи з того, що ця модель належить до експоненційного класу, розподіл кожної окремої несправності має вигляд:

$f_X(x)=b\exp(-bx)$, тоді

$$\lambda(t) = Nb\exp(-bt) = Nf_X(t) -$$

це залежність між функцією інтенсивності відмов та функцією щільності ймовірності для окремої несправності.

Виходячи з припущень про те, що **f_i** – незалежні пуасонівські випадкові величини з математичним сподіванням $= \mu(t_i) - \mu(t_{i-1})$, сукупна щільність для **f_i**-их має вигляд:

$$\prod_{i=1}^{n}\frac{[\mu(t_i)-\mu(t_{i-1})]^{f_i}\exp(\mu(t_i)-\mu(t_{i-1}))}{f_i!})$$

Оцінювання моделі

За методом максимальної вірогідності Гоель та Окумото отримали наступні оцінки параметрів своєї моделі:

$$\tilde{N}=\frac{\sum_{i=1}^{n} f_i}{(1-\exp(-\tilde{b}t_n))} \text{ та } \frac{t_n\exp(-\tilde{b}t_n)\sum_{i=1}^{n} f_i}{(1-\exp(-\tilde{b}t_n))}=\sum_{i=1}^{n}\frac{f_i(t_i\exp(-\tilde{b}t_i)-t_{i-1}\exp(-\tilde{b}t_{i-1})}{\exp(-\tilde{b}t_{i-1})-\exp(-\tilde{b}t_i)}$$

Друге рівняння розв'язується чисельними методами й отримується оцінка для N, яку підставляється в перше, після чого отримуються всі необхідні параметри для цієї моделі.

Надійність

Пропонуємо розглянути додатково функцію надійності цієї моделі:

$$R(\Delta t|t_{i-1})=\exp(-[N\exp(-bt_{i-1})][1-\exp(-b\Delta t)])$$

$$z(\Delta t|t_{i-1})=Nb\exp(-bt_{i-1})\exp(-b\Delta t)$$

Переваги

Простота застосування, легкість обчислення параметрів. За допомогою функції надійності можна побачити, що при $t\to\infty$ надійність ПЗ $R(t)\to 1$, отже ПЗ може стати надійним.

Недоліки

Функція інтенсивності несправностей строго спадає, проте це не відповідає результатам, отриманим на практиці - вона може, як спадати так і тимчасово зростати.

Модель Шнайдевінда

В основі даної моделі лежить наступна ідея: не всі дані, щодо кількості несправностей в інтервал часу, мають однаковий вплив, можливо за допомогою пізніше отриманих даних можна отримати точніші прогнози надійності ніж при використанні більш ранніх даних. Для реалізації та аналізу даної ідеї Шнайдевіндом було запропоновано три варіанти моделі [6]. В моделі використовуються дані про кількість несправностей у проміжок часу, виходячи з того, що всі часові інтервали рівні за величиною. Нехай в нас є *n* часових інтервалів, на яких було зібрано відповідну інформацію, тоді три варіанти моделі виглядають наступним чином:

Модель 1 використовуються дані з усіх n інтервалів. Це відображує рівнозначність даних з усіх проміжків;

Модель 2 ігноруються дані з проміжків від 1 до s-1 включно, тобто до аналізу беруться лише дані з s до n тестового інтервалу. Це відображує те, що більш ранні данні дають менший вклад (якщо взагалі впливають) у прогнозування надійності;

Модель 3 в цьому варіанті використовується кумулятивна кількість несправностей з проміжків від 1 до s-1 як базова, а дані з інтервалів від s до n вважаються допоміжними. Це значить, що комбінація даних з s-1 перших інтервалів відображає процес виявлення несправностей на більш пізніх стадіях.

Припущення моделі:

Враховуючи стандартні, модель має наступні припущення:

1. Кумулятивна кількість відмов на момент t, M(t) відповідає пуасонівському процесу з функцією математичного сподівання $\mu(t)$. Мат. сподівання таке, що очікувана кількість несправностей для будь-якого тестового часового інтервалу пропорційна кількості несправностей, що ще залишилися у системі на цей час. Також вважаємо, що $\mu(t)$ – обмежена, не спадаюча функція часу: $\lim_{t\to\infty}\mu(t)=\alpha/\beta<\infty$, для констант $\alpha,\beta>0$ Отже, ця модель належить до категорії скінчених функцій.
2. Вважається, що функція інтенсивності відмов характеризується експоненційним розподілом $\lambda(t)=\alpha\exp(-\beta t)$.

   Вочевидь, що більші β відповідають меншим значенням функції інтенсивності відмов. Більш того, параметр α - початкове значення $\lambda(t)$.
3. Кількості несправностей $(f_1,f_2,\ldots,f_n)$, виявлених на кожному з часових інтервалах незалежні.
4. Частота виправлення несправностей пропорційна кількості нестравностей, що треба виправити.
5. Часові інтервали, під час яких досліджується ПЗ, рівні за величиною, тобто $t_i=il$, для $i=1,\ldots,n$, а l – деяка додатна константа (не втрачаючи загальності можна вважати, що $l=1$, тоді $t_i=i$)

Для моделі Шнайдевінда використовують такі **набори даних**:

$\mathbf{f_i}$ - кількість несправностей на кожному з досліджуваних інтервалів;

Формальний вигляд моделі

Враховуючі припущення моделі й її опис, можна записати, що $\mu(t) = \frac{\alpha}{\beta}(1 - \exp(-\beta t))$,

отже $\mu(t_i) = \frac{\alpha}{\beta}(1 - \exp(-\beta i))$, тоді очікувана кількість несправностей в i-ий період часу:

$$m_i = \mu(t_i) - \mu(t_{i-1}) = \frac{\alpha}{\beta}[\exp(-\beta(i-1)) - \exp(-\beta i)]$$

Виходячи з припущень про те, що **f**$_i$ – незалежні пуасонівські випадкові величини та те, що існує декілька різновидів моделі, сукупна щільність для **f**$_i$-их має вигляд:

$$\frac{M_{s-1}^{F_{s-1}} \exp(-M_{s-1})}{F_{s-1}} \prod_{i=s}^{n} \frac{m_i^{f_i} \exp(-m_i)}{f_i!},$$

де s – деяке ціле число від 1 до n , M_{s-1} – кумулятивне мат. сподівання кількості несправностей в системі до інтервалу s –1 включно, F_{s-1} - кумулятивна кількість несправностей, виявлених за інтервали часу до s –1 включно.

Оцінювання моделі

За методом максимальної вірогідності Шнайдевінд отримав наступні оцінки параметрів своєї моделі:

Модель 1

$$\tilde{\alpha} = \frac{\tilde{\beta} F_n}{1 - \exp(-\tilde{\beta} n)} \quad \text{та} \quad \frac{1}{\exp(\tilde{\beta}) - 1} - \frac{n}{\exp(\tilde{\beta} n) - 1} = \sum_{k=0}^{n-1} k \frac{f_{k+1}}{F_n}, \text{ де } F_n = \sum_{i=1}^{n} f_i$$

Модель 2

$$\tilde{\alpha} = \frac{\tilde{\beta} F_{s,n}}{1 - \exp(-\tilde{\beta}(n - s + 1))} \quad \text{та} \quad \frac{1}{\exp(\tilde{\beta}) - 1} - \frac{n - s + 1}{\exp(\tilde{\beta}(n - s + 1)) - 1} = \sum_{k=0}^{n-s} k \frac{f_{k+s}}{F_{s,n}}, \text{ де } F_{s,n} = \sum_{i=s}^{n} f_i$$

вочевидь, що при s =1 моделі 1 та 2 еквівалентні.

Модель 3

$$\tilde{\alpha} = \frac{\tilde{\beta} F_n}{1 - \exp(-\tilde{\beta} n)} \quad \text{та} \quad \frac{(s-1)F_{s-1}}{\exp(\tilde{\beta}(s-1)) - 1} + \frac{F_{s,n}}{\exp(\tilde{\beta}) - 1} - \frac{nF_n}{\exp(\tilde{\beta} n) - 1} = \sum_{k=0}^{n-s} (s + k - 1) f_{k+s},$$

де $F_{s-1} = \sum_{i=1}^{s-1} f_i$

аналогічно при s =1 моделі 1 та 3 еквівалентні.

Робота [6] присвячена вибору параметра s.

Надійність

Функція надійності цієї моделі має вигляд:

$$R(\Delta t | t_{i-1}) = \exp(-[\alpha \exp(-\beta t_{i-1})][1 - \exp(-\beta \Delta t)])$$

$$z(\Delta t | t_{i-1}) = \alpha \exp(-\beta t_{i-1}) \exp(-\beta \Delta t)$$

Переваги

Різниця у впливі ранніх та більш пізніх даних. За допомогою функції надійності можна побачити, що при $t \to \infty$ надійність ПЗ $R(t) \to 1$, отже ПЗ може стати надійним.

Недоліки

Складність обчислення параметра s.

Узагальнення 1

Слід відмітити, що якщо у **моделі негомогенного пуасонівського процесу** всі досліджувані часові інтервали взяти однакової довжини та параметр $N=\alpha/\beta$, то отримаємо перший варіант моделі Шнайдевінда.

Базова модель Муси

Модель Муси – найбільш поширена серед моделей зростання надійності. Вона була розроблена Дж. Мусою в лабораторії AT&T Bell, який розробив ще цілий ряд моделей [8,9,10,11]. Це одна з перших моделей, в якій використовується фактичний час роботи ПЗ, а не календарний. Муса вважав, що фактичний час краще відображує процес ніж календарний. Проте в моделі передбачена можливість переведення фактичного часу до календарного [9].

Припущення моделі:

Враховуючи базові, модель має наступні припущення:

1. Кумулятивна кількість відмов на момент t, M(t) відповідає пуасонівському процесу з функцією математичного сподівання
$$\mu(t) = \beta_0(1 - \exp(-\beta_1 t)).$$
Мат. сподівання таке, що очікувана кількість несправностей для будь-якого тестового часового інтервалу пропорційна кількості несправностей, що ще залишилися у системі на цей час. Так як $\lim_{t\to\infty}\mu(t) = \lim_{t\to\infty}(\beta_0(1-\exp(-\beta_1 t))) = \beta_0 < \infty$, ця модель належить до категорії скінчених функцій. Параметр β_0- загальна кількість несправностей. що буде виявлено.
2. Вважається, що час між відмовами характеризується кусково-експоненційним розподілом, тобто z(t) для окремої несправності є константою. Отже, модель належить до класу експоненціальних моделей.

Зауваження: тут приведені припущення для використання фактичного часу, з додатковими умовами, що виникають в разі використання календарного часу можна ознайомитися в [9].

Для базової моделі Муси використовують такі набори даних:

або фактичні значення часу, коли відбулися відмови - ПЗ - $\mathbf{t_1, t_2, \dots t_n}$

або час між відмовами $\mathbf{x_1, x_2, \dots x_n}$. Вони пов'язані наступним співвідношенням: $\mathbf{x_i = t_i - t_{i-1}}$.

Формальний вигляд моделі

Виходячи з того, що
$$\mu(t) = \beta_0(1 - \exp(-\beta_1 t)),$$
функція інтенсивності відмов матиме вигляд:
$$\lambda(t) = \beta_0\beta_1\exp(-\beta_1 t)).$$
Вочевидь, що чим більше β_1 тим швидше спадає функція інтенсивності і навпаки.

Оцінювання моделі

Нехай спостерігалося **n** відмов в моменти часу $\mathbf{t_1, t_2, \dots t_n}$, та з моменту $\mathbf{t_n}$ пройшов час **x** без відмов ($\mathbf{t_n+x}$ – загальний час спостереження ПЗ). Тоді використовуючи припущення моделі можна отримати функцію подібності для цього класу:
$$L(\beta_0,\beta_1) = \beta_0^n\beta_1^n[\prod_{i=s}^{n}\exp(-\beta_1 t_i)]\exp(-\beta_0[1-\exp(-\beta_1(t_n+x))])$$
Отже, за методом максимальної вірогідності Муса отримав оцінки параметрів для своєї моделі шляхом розв'язку системи рівнянь.
$$\tilde{\beta}_0 = \frac{n}{1-\exp(-\tilde{\beta}_1(t_n+x))}$$
та $\dfrac{n}{\tilde{\beta}_1} - \dfrac{n(t_n+x)}{\exp(-\tilde{\beta}_1(t_n+x))-1} = \sum_{k=0}^{n} t_i$

Надійність

Функція надійності цієї моделі має наступний вигляд:
$$R(\Delta t | t_{i-1}) = \exp(-[\beta_0\exp(-\beta_1 t_{i-1})][1-\exp(-\beta_1\Delta t)]$$
$$z(\Delta t | t_{i-1}) = \beta_0\beta_1\exp(-\beta_1 t_{i-1})\exp(-\beta_1\Delta t)$$

Переваги

Простота застосування, легкість обчислення параметрів, використання фактичного часу роботи. За допомогою функції надійності можна побачити, що при $t \to \infty$ надійність ПЗ $R(t) \to 1$, отже ПЗ може стати надійним.

Недоліки

Функція інтенсивності несправностей строго спадає, проте це не відповідає результатам, отриманим на практиці.

S-подібна модель зростання надійності

Ця модель є представником класу гамма-розподілів. Тобто розподіл кожної несправності – гамма. Проте кількість відмов за часовий інтервал характеризується пуасонівським розподілом. Обрана форма кривої зростання надійності (S-подібна) пояснюється в такий спосіб: спочатку йде крива навчання, потім тестова команда ознайомлюється з ПЗ і крива починає швидко зростати поки нарешті не переходить в третю фазу - так, що несправності все важче виявити. Вона була запропонована Ямадою, Обою та Осакі [12,13].

Припущення моделі:
Враховуючи базові, модель має наступні припущення:

1. Кумулятивна кількість відмов на момент t, M(t) відповідає пуасонівському процесу з функцією математичного сподівання
$$\mu(t) = \alpha(1-(1+\beta t)\exp(-\beta t)).$$
Це обмежена незростаюча функція часу, так як $\lim_{t\to\infty}\mu(t) = \alpha < \infty$, отже ця модель належить до категорії скінчених функцій.
2. Час між відмовами (i-1) та (i) залежить від часу до (i-1)-ої відмови.
3. Несправність видаляється відразу ж після її виявлення, й при процесі корекції не вносяться нові несправності.

Для S-подібної моделі зростання надійності використовують такі набори даних:
або значення часу, коли відбулися відмови - ПЗ **- t_1, t_2, … t_n** або $\mathbf{f_i}$ - кількість несправностей на кожному з досліджуваних інтервалів, зважаючи на те, що $\mathbf{l_i}$ – довжини цих періодів.

Формальний вигляд моделі

Нехай є наступне розбиття часу (тестові інтервали) $t_0^* = 0 < t_1^* < t_2^* < ... < t_n^*$. Позначимо його через T^*. f_1, f_2,…,f_n – кількості несправностей виявлених на кожному з часових інтервалів, тобто f_i – кількість несправностей, які були виявлені на інтервалі довжиною $l_i = t_i^* - t_{i-1}^*$.

З базових припущень маємо:
$$\mu(t_i^*) - \mu(t_{i-1}^*) = \alpha(1-(1+\beta t_i^*)\exp(-\beta t_i^*)) - \alpha(1-(1+\beta t_{i-1}^*)\exp(-\beta t_{i-1}^*)) =$$
$$= \alpha[(1+\beta t_{i-1}^*)\exp(-\beta t_{i-1}^*) - (1+\beta t_i^*)\exp(-\beta t_i^*)]$$
тоді функція інтенсивності матиме вигляд:
$$\lambda(t) = \alpha\beta^2 t\exp(-\beta t))$$ - це гамма-розподіл.

Через вигляд функції $\mu(t)$ модель і отримала назву S-подібної. В ній параметр β відповідає за усунення несправностей (їх виявлення та корекцію). Так як $\lim_{t\to\infty}\mu(t) = \alpha < \infty$, то за класифікацією ця модель належить до категорії скінчених функцій, параметр α - є загальною кількістю несправностей у системі.

Оцінювання моделі

Виходячи з припущень моделі, сукупна щільність для $\mathbf{f_i}$-их:
$$\prod_{i=1}^{n}\frac{[\mu(t_i^*) - \mu(t_{i-1}^*)]^{f_i}\exp(\mu(t_i^*) - \mu(t_{i-1}^*))}{f_i!}.$$
За методом максимальної вірогідності були отримані наступні оцінки параметрів для S-подібної моделі зростання надійності:
$$\sum_{i=1}^{n} f_i = \tilde{\alpha}(1-(1+\tilde{\beta}t_n^*)\exp(-\tilde{\beta}t_n^*))$$ та
$$\tilde{\alpha}(t_n^*)^2\exp(-\tilde{\beta}t_n^*) = \sum_{i=1}^{n}\frac{(\sum_{j=1}^{i} f_j - \sum_{j=1}^{i-1} f_j)((t_i^*)^2\exp(-\tilde{\beta}t_i^*) - (t_{i-1}^*)^2\exp(-\tilde{\beta}t_{i-1}^*))}{((1+\tilde{\beta}t_{i-1})\exp(-\tilde{\beta}t_{i-1}) - (1+\tilde{\beta}t_i)\exp(-\tilde{\beta}t_i))}.$$

Ці рівняння розв’язуються чисельними методами. Якщо у нас є $\mathbf{t_i}$, то $t_i = t_i^*$, а якщо $\mathbf{f_i}$**,** асоційовані з довжиною періоду, тоді $t_i^* = \sum_{j=1}^{i} l_j$.

Надійність

Функція надійності цієї моделі має наступний вигляд:
$$R(\Delta t|t_{i-1}) = \exp(-\alpha\exp(-\beta t_{i-1})[1+\beta t_{i-1} - \exp(-\beta\Delta t) - \beta(t_{i-1}+\Delta t)\exp(-\beta\Delta t)])$$
$$z(\Delta t|t_{i-1}) = \alpha\beta^2(\Delta t + t_{i-1})\exp(-\beta t_{i-1})\exp(-\beta\Delta t).$$

Переваги

S-подібна форма функції інтенсивності відповідає практиці проведення тестування ПЗ [12]. За допомогою функції надійності можна побачити, що при $t\to\infty$ надійність ПЗ $R(t)\to 1$, отже ПЗ може стати надійним.

Недоліки

Складність обчислення параметрів. Припущення 3 цієї моделі звужує можливості її застосування тому, що на практиці велика кількість несправностей вноситься саме на фазі корекції програмного забезпечення.

Узагальнення моделі негомогенного пуасонівського процесу

Ми розглянули основні відомі моделі, що базуються на негомогенному пуасонівському процесі - моделі Гоеля-Окумото, Шнайдевінда, Муси та S-подібну модель. Автори проаналізувати якість роботи розглянутих моделей на практиці на базі реальних даних про кількість виявлених несправностей та їхній розподіл у часі за допомогою програми CASRE 3.0, що була надана фірмою Open Channel Software (рис.1 та рис. 2).

Перший графік (Рис.1) ілюструє можливості моделі Муси в оцінюванні інтенсивності. Як бачимо, із зростанням часу модель дає кращі результати, але на початку розглядаємого часового проміжку прогноз не є якісним. Аналогічні графіки були отримані для моделей Гоеля-Окумото та Шнайдевінда. Модель починає давати більш якісні прогнози після 13,8 годин роботи ПЗ. Можна бачити, що формальний вигляд моделей Гоеля-Окумото, Муси та Шнайдевінда (варіант1) є ідентичним, що знаходить своє відображення на ілюструючих графіках – графіки даних моделей співпадають.

Таким чином, інтенсивність виявлення несправностей спадає експоненційно, проте по проходженню певного часового проміжка. Завдання полягає в зменшенні цього часу.

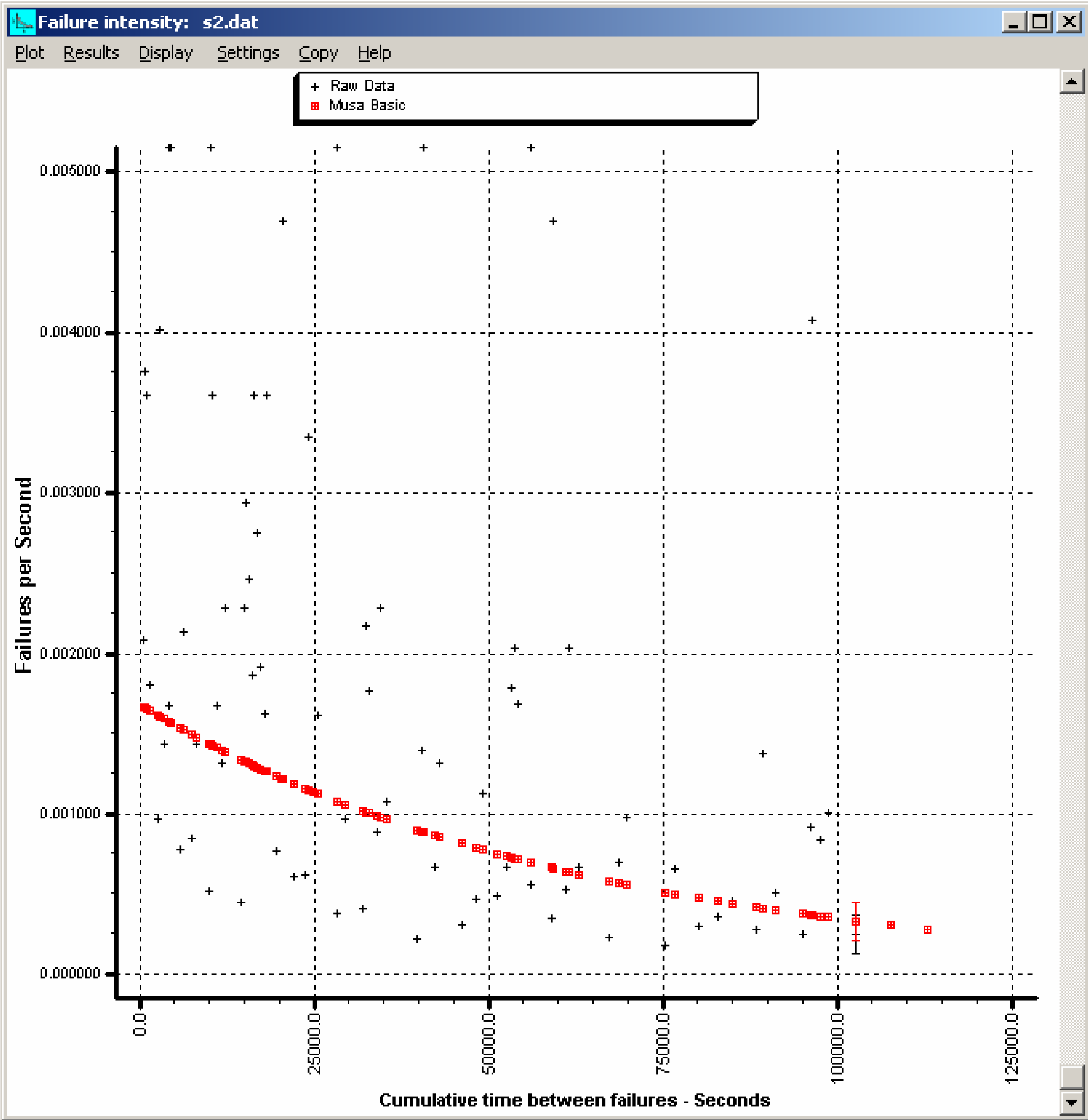

Рис. 1. Модель Муси

Другий графік (Рис.2) ілюструє можливості моделей Гоеля-Окумото, Шнайдевінда та S-подібної моделі зростання надійності в оцінюванні інтенсивності. Для моделі Муси график співпадає з графіком для моделі Гоеля-Окумото, проте через технічні обмеження програми CASRE 3.0 не вдалося розташувати всі бажані чотири результати на одному графіку.

Результати, що тут розміщені, дають змогу порівняти якість прогнозів. Вочевидь, що із зростанням часу всі моделі спадають експоненційно і мають схожі результати. Проте на початку розглянутого часового проміжку прогнози доволі відрізняються.

Моделі Гоеля-Окумото, Шнайдевінда та Муси не відображують дійсну картину інтенсивності виявлення несправностей протягом майже 13 годин. Схожі дані були отримані і при аналізі першого графіка.

Прогноз, отриманий за допомогою S-подібної моделі більш якісний. Вона починає давати якісні прогнози приблизно після 7 годин роботи ПЗ. Більш того, форма кривої S-подібної моделі в перші 7 годин схожа на форму кривої розподілу інтенсивності виявлення несправностей. Різниця полягає в значеннях екстремумів, часу іхнього досягнення та деяких інших деталях.

Зауважимо, що часовий проміжок, необхідний для початку роботи моделі став меншим.

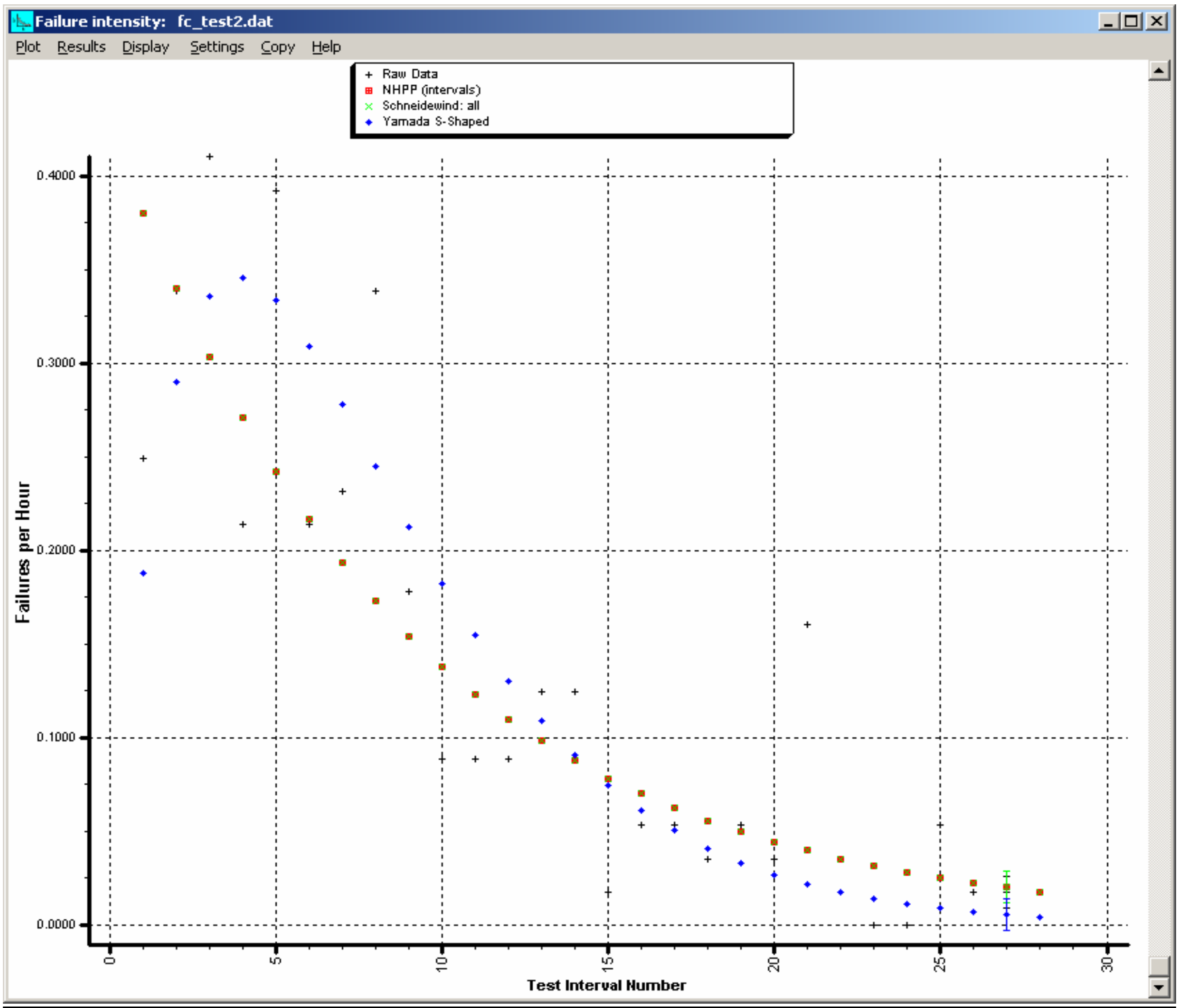

Рис. 2. Модель Гоеля-Окумото, Шнайдевінда та та S-подібна модель зростання надійності

Скінчена модель не гомогенного пуасонівсього процесу.

В ході проведення дослідження авторами було запропоноване узагальнення та розширення моделей, що базуються на пуасонівському процесі.

За базові були взяті наступні моделі: модель Гоеля-Окумото, модель Шнайдевінда, базова модель Муси, та S-подібна модель зростання надійності.

Як показано вище, фактично моделі Гоеля-Окумото, Шнайдевінда та Муси відрізняються лише методами оцінки параметрів, а сама форма моделі для них ідентична. Сукупний недоліком цих моделей є форма кривої інтенсивності виявлення несправностей, вона є експоненційною та строго спадає при $t > 0$. На практиці ж було показано, що це не відповідає дійсності та не достатньо відображує процес тестування [13]. З метою корекції цього недоліка автори взяли ще одну базову модель - S-подібну модель зростання надійності.

Функція інтенсивності виявлення несправностей цієї моделі : $\lambda(t) = \alpha\beta^2 t \exp(-\beta t))$

Вона відрізняється від функцій інтенсивності моделей Муси, Гоеля-Окумото та Шнайдевінда введенням параметру t, який дає змогу змінити форму кривої таким чином, що спочатку інтенсивність зростає, а потім строго спадає. Проте зважаючи на наявність експоненти ця зміна не є довготривалою.

З метою більш точно відтворювати форму кривої інтенсивності та підвищення гнучкості цієї моделі автори пропонують розглядати наступну форму кривої інтенсивності: $\lambda(t) = \alpha\beta^{n+1}t^n \exp(-\beta t))$,

в якій додатковий параметр n - характеризує складність та розміри проекта.

У невеличких проектах для яких розробник є і тестером функціонують моделі Муси, Гоеля-Окумото та Шнайдевінда, отже параметр n=0; для більш складних проектів з відокремленими функціями програміста та тестера параметр n=1 і та. ін. Таким чином, параметр n зростає зі складністю проекта і допомагає відобразити наступну практичну ситуацію тестування проекта:

- спочатку група тесторів не знайома з проектом, відповідно вони знаходять дуже незначну кількість помилок – інтенсивність виявлення помилок близька до нуля;
- з накопиченням досвіду інтенсивність виявлення помилок зростає;
- нарешті в системі залишається невелика кількість несправностей і виявлення кожної з них потребує дуже значного часу – інтенсивність виявлення помилок спадає.

Всі базові моделі належать до категорії скінчених функцій, результуюча узагальнена модель також належить до категорії скінчених функцій. Функції надійності цих моделей $R(t) \to 1$ при $t \to \infty$, отже ПЗ може стати надійним. Ця властивість була збережена і в результуючій моделі.

Враховуючи базові, узагальнена модель має наступне припущення - кумулятивна кількість відмов на момент t, M(t) має відповідати пуасонівському процесу з функцією математичного сподівання $\mu(t)$. Математичне сподівання таке, що очікувана кількість несправностей, які буде виявлено за проміжок часу $t+\Delta t$ пропорційна кількості несправностей, що ще залишилися у системі на момент t.

Представимо зв'язки між скінченими моделями не гомогенного пуасонівського процесу, які були розглянуті вище та результуючу узагальнену модель.

Таблиця 1

Зв'язок між скінченими моделями не гомогенного пуасонівського процесу

Модель	$\lambda(t)$	$\mu(t)$
Модель Гоеля-Окумото	$\lambda(t) = Nb\exp(-bt)$	$\mu(t) = N(1-\exp(-bt))$
Модель Шнайдевінда	$\lambda(t) = \alpha\exp(-\beta t)$	$\mu(t) = \frac{\alpha}{\beta}(1-\exp(-\beta t))$
Базова модель Муси	$\lambda(t) = \beta_0\beta_1\exp(-\beta_1 t))$	$\mu(t) = \beta_0(1-\exp(-\beta_1 t))$
S-подібна модель зростання надійності	$\lambda(t) = \alpha\beta^2 t\exp(-\beta t))$	$\mu(t) = \alpha(1-(1+\beta t)\exp(-\beta t))$
Узагальнена та розширена скінчена модель не гомогенного пуасонівського процесу	$\lambda(t) = \alpha\beta^{n+1}t^n\exp(-\beta t))$	$\mu(t) = a(1-\sum_{i=0}^{n}\frac{n!b^{n-i}}{(n-i)!}t^{n-i}\exp(-bt)))$

Таким чином, авторами була отримана узагальнена скінчена модель не гомогенного пуасонівського процесу, яка наслідує своїм базовим моделям, її функція надійності $R(t) \to 1$ при $t \to \infty$, отже ПЗ може стати надійним. Функція інтенсивності виявлення помилок має форму, що збігається з результатами, отриманими на практиці.

Параметри α та β обчислюються за методом максимальної вірогідності, в галузі оцінювання параметра n авторами проводиться подальша робота.

Висновки

Авторами даної статті було проведено дослідження моделей оцінювання надійності програмного забезпечення, що базуються на не гомогенному пуасонівському процесі. В результаті даного дослідження були виявлені певні недоліки моделей Муси, Гоеля-Окумото, Шнайдевінда та S- подібної моделі зростання надійності, пов'язані з відмінностями від практики проведення тестування та неякісним оцінюванням інтенсивності виявлення несправностей на початкових етапах.

З метою усунення цих недоліків, а в особливості зменшення часового проміжку, необхідного для початку роботи моделі, проведена розробка узагальненої та розширеної моделі не гомогенного пуасонівського

процесу, що дозволило ускладнити форму кривої інтенсивності виникнення помилок з урахуванням процесів навчання, що відповідає практиці [12].
Авторами ведуться подальші роботи в галузі оцінки параметрів.

Література

1. F. Saglietti,"Softwarezuverlässigkeit - 20 Jahre später"Mitteilungen der Fachgruppe "Fehlertolerierende Rechensysteme" der Gesellschaft für Informatik, Nr. 30, 2002
2. S.S. Gokhale, P.N.Marinos, K.S.Trivedi, “Important Milestones in Software Reliability Modeling”, In Proc. of Software Engineering and Knowledge Engineering, 1996
3. Handbook of Software Reliability Engineering, Editor M. Lyu, McGraw Hill, , 1996.
4. J.C. Laprie. Dependability - its attributes, impairments and means. In Predictably Dependable Computing Systems, chapter I, Basic Concepts, pages 3-24, Springer Verlag, 1995.
5. A.L.Goel, K.Okumoto, “Time-Dependent Error-Detection Rate Model for Software and other Performance Measures”, IEEE Transactions on Reliability, vol. R-28, no. 3, August 1979, pp. 206-211.
6. N.F. Schneidewind, “Analysis of Error Process in Computer Software”, Sigplan Note, vol. 10, no.6, 1975, pp.337-346.
7. N.F. Schneidewind, “Software Reliability Model with Optimal Selection of Failure Data”, IEEE Transactions on Software Engineering, vol. 19, no.11, November 1993, pp. 1095-1104.
8. J.D. Musa, “A theory of software reliability and its application.” IEEE Transactions on Software Engineering SE-1(3), 312–327, 1975
9. J.D. Musa, A. Iannino, and K. Okumoto, “Software Reliability: Measurement, Prediction, Application”, McGraw-Hill, New York, 1987.
10. J.D. Musa, “Software Reliability Engineering”, McGraw-Hill, New York, 1998.
11. J.D. Musa, K. Okumoto, “A logarithmic poisson execution time model for software reliability measurement.” In Proc. Seventh International Conference on Software Engineering, Orlando, Florida, pp. 230–238, 1984
12. M. Ohba , “Software Reliability Analysis Models,” IBM J. Res. Develop, Vol. 28, No. 4, July 1984, pp.428-243.
13. S.Yamada, M.Ohba, S.Osaki, “S-Shaped reliability Grows Modeling for Software Error Detection”, IEEE Transactions on Reliability, vol. R-32, no.5, December 1983, pp. 475-478.