

УДК 004.891.3: 004.3

ОЦІНЮВАННЯ ІНФОРМАТИВНОСТІ ТА ЕФЕКТИВНОСТІ МЕТРИК НА ЕТАПІ ПРОЕКТУВАННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

Т.О.ГОВОРУЩЕНКО, А.В.БАЧИНСЬКИЙ, Є.В.ПИТЛИК

Хмельницький національний університет

У статті досліджено інформативність та ефективність метрик, використовуваних на етапі проектування програмного забезпечення. Проведене дослідження показало, що ефективність та інформативність метрики не залежать від величини її діапазону.

In this paper the informativeness and the efficiency of software design stage metrics are researched. This research represents, that the efficiency and the informativeness of metrics don't depend on the size of metric's range.

Ключові слова: Метрики складності та якості програмного забезпечення, інформативність метрик, релевантність метрик, достовірність метрик, значущість метрик, ефективність метрик.

Вступ. Якість програмного забезпечення (ПЗ) визначається якістю методів та інструментальних засобів, які застосовувались для забезпечення всього їх життєвого циклу. На практиці важливо оцінювати якість програм не лише в завершеному вигляді, але й в процесі їх проектування і розроблення. Оцінки складності та якості проекту є важливими для одержання об'єктивних оцінок щодо проекту. Вказані показники не можуть бути обчислені на ранніх стадіях роботи над проектом, оскільки вимагають, як мінімум, детального проектування. Однак ці показники важливі для одержання прогностичних оцінок тривалості та вартості проекту, оскільки безпосередньо визначають його трудомісткість.

Одним з основних засобів аналізу і оцінювання складності та якості ПЗ є метричний аналіз. Метрика ПЗ - це міра, яка дозволяє одержати числове значення деякої властивості ПЗ або його специфікацій. Складність обґрунтування вибору та інтерпретації метрик в процедурах прийняття виробничих рішень та ігнорування етапів життєвого циклу ПЗ не дозволяють повноцінно

використовувати метрики для підвищення якості ПЗ та зменшення його складності.

Отже, актуальною задачею є виділення метрик, які можливо використати вже на етапі проектування, а також визначення їх інформативності та ефективності для проекту ПЗ.

1. Метрики складності та якості програмного забезпечення на етапі проектування. На етапі проектування важливо закласти цілу низку вимог по якості: вимоги до структури програмного забезпечення; вимоги до навігації по ПЗ; вимоги до дизайну інтерфейсів користувача; вимоги до мультимедіа-компонентів ПЗ; вимоги по зручності (usability); технічні вимоги. Отже, на етапі проектування формується відповідь на питання "Яким чином програмна система буде реалізовувати висунуті до неї вимоги?".

Інформаційні потоки етапу проектування ПЗ [1]: вимоги до ПЗ, представлені інформаційною, функційною та поведінковою моделями аналізу. Інформаційна модель описує інформацію, які повинна обробляти ПЗ на думку замовника. Функційна модель визначає перелік функцій обробки інформації та перелік модулів програмної системи. Поведінкова модель фіксує бажану динаміку системи (режими її роботи). На виході етапу проектування - розроблення даних, розроблення архітектури та процедурне розроблення ПЗ.

Виходячи з інформації, доступної на етапі проектування, і аналізу метрик складності та якості програмного забезпечення [2, 3], оберемо метрики ПЗ, які можна використовувати вже на етапі проектування (таблиця 1).

Таблиця 1. Метрики складності та якості програмного забезпечення, використовувані на етапі проектування

№ п/п	Метрики складності ПЗ	Метрики якості ПЗ
(1)	(2)	(3)
1	Метрика Чепіна	Метрика зв'язності
2	Метрика Джилба (абсолютна)	Метрика зчеплення
3	Метрика Мак-Клур (загальна складність ПЗ)	Метрика звертання до глобальних змінних
4	Метрика Кафура (інформаційна складність модуля)	Час модифікації моделей (в роб. днях)

(1)	(2)	(3)
5	Очікувана LOC-оцінка	Загальна кількість знайдених помилок при інспектуванні моделей та прототипів модулів
6	Метрика Холстеда (складність програми)	Прогнозований загальний час розроблення ПЗ (в роб. днях)
7	Метрика Маккейба (цикломатична складність ПЗ)	Час виконання робіт процесу проектування ПЗ (в роб. днях)
8	Метрика Джилба (відносна)	Очікувана вартість розроблення ПЗ (в грн)
9	Прогнозована кількість операторів програми	Прогнозована вартість перевірки якості ПЗ
10	Прогнозована оцінка складності інтерфейсів ПЗ	Прогнозована продуктивність розроблення ПЗ (в хвилинах на один рядок коду)
11		Прогнозовані витрати на реалізацію програмного коду (в грн)
12		Прогнозований функційний розмір
13		Прогнозована оцінка трудовитрат за моделлю Боема (в людиномісяцях)
14		Прогнозована оцінка тривалості проекту за моделлю Боема (в місяцях)

2. Оцінювання інформативності та ефективності метрик

ПЗ. Оцінимо інформативність, яку містять метрики складності та якості ПЗ. Критерії оцінювання інформативності [4]: 1) релевантність - наявність зв'язку з проблемою (відповідність інтересам та проблемі) та здатність інформації внести ясність у процес розуміння проблеми; 2) достовірність - наскільки представлений опис відповідає дійсності; 3) значущість - розуміння інформації, повнота висвітлення предмету інтересу, своєчасність інформації та її достатність для прийняття рішень, а також кількість інформації:

$$I = H(\text{априорна}) - H(\text{апостеріорна}),$$

де $H(\text{априорна})$ - ентропія системи (показник незнання системи) до експерименту,

$H(\text{апостеріорна})$ - ентропія системи після експерименту, тоді, якщо $I > 0$, то в результаті експерименту одержано нову інформацію, показник незнання системи зменшився.

Щодо релевантності метрик складності і якості ПЗ, то рівень релевантності дорівнює 1, оскільки саме метрики якості дають уявлення про складність та якість проекту і ПЗ. Достовірність метрик дорівнює 1, оскільки використовуються лише класичні метрики складності та якості, описані в галузевій літературі.

Оскільки релевантність і достовірність метрик якості ПЗ є сталими величинами, то основним критерієм оцінювання інформації, який впливатиме на оцінку ефективності метрик якості ПЗ, є саме значущість інформації (інформативність) метрик.

Сформуємо підсумковий критеріальний показник ефективності для кожної метрики за адитивним критерієм.

Адитивний критерій A формується шляхом ділення суми добутків часткових показників ефекту l_i на коефіцієнти значущості i -го показника g_i на кількість показників ефекту n :

$$A = (1/n) \cdot \sum_{i=1}^n (l_i \cdot g_i) \quad [5].$$

Значущість, а відтак і ефективність, кожної метрики якості ПЗ етапу проектування залежить від 4-х показників ефекту ($n=4$): розуміння інформації, повнота висвітлення предмету інтересу, своєчасність інформації та її достатність для прийняття рішень, кількість інформації. Кожен показник ефекту розглядатимемо для спрощення розрахунків як значення з інтервалу $[0, 1]$, де 0 - мінімальне значення показника, 1 - максимальне значення. Значущості показників

ефекту для кожної метрики рівні, причому $\sum_{i=1}^n g_i = 1$, отже,

$$g_i = \frac{1}{4} = 0,25.$$

Також необхідно розрахувати 4 часткових показники ефекту l_i для кожної метрики складності та якості ПЗ і відповідно ефективність кожної метрики за адитивним критерієм.

Показник "розуміння інформації" l_1 дорівнює 1 для кожної з метрик, оскільки метрики складності та якості розраховані вірно, з повним розумінням.

Показник "повнота висвітлення предмету інтересу" l_2 залежить від повноти висвітлення метрикою оброблюваної інформації, функцій обробки інформації та переліку модулів, режимів роботи ПЗ. Якщо метрика в повній мірі висвітлює всі 3 вищезазначені характеристики, то показник "повнота висвітлення предмету інтересу" дорівнює $l_2 = 1$. Тоді, якщо метрика в повній мірі висвітлює лише одну з характеристик, то показник "повнота висвітлення предмету інтересу" для неї становить $l_{2_1} = \frac{1}{3} = 0,333$. Характеристика "функції обробки інформації та перелік модулів" складається відповідно з двох частин, тому висвітлення метрикою однієї з частин характеристики складає значення показника $l_{2_{1/2}} = \frac{0,333}{2} = 0,1665$. Відповідно, показник l_2 дорівнює сумі вищеописаних часткових значень залежно від повноти висвітлення метрикою оброблюваної інформації.

Показник "своєчасність інформації та її достатність для прийняття рішень" l_3 залежить від достатності інформації про оброблювану інформацію, функції обробки інформації та перелік модулів, а також про режими роботи ПЗ. Якщо метрика залежить від усіх 3-х вищезазначених характеристик, то показник "своєчасність інформації та її достатність для прийняття рішень" дорівнює $l_3 = 1$. Тоді, якщо метрика залежить лише від однієї характеристики, то показник "своєчасність інформації та її достатність для прийняття рішень" для неї становить $l_{3_1} = \frac{1}{3} = 0,333$. Характеристика "функції обробки інформації та перелік модулів" складається з двох частин, тому залежність метрики від однієї з частин характеристики складає значення показника $l_{3_{1/2}} = \frac{0,333}{2} = 0,1665$. Відповідно, показник l_3 дорівнює сумі вищеописаних часткових значень залежно від висвітлення метрикою оброблюваної інформації.

Для розрахунку показника "кількість інформації" l_4 вважатимемо, що після опрацювання всіх метрик якості є повна інформація про якість проекту та ПЗ, тому кількість інформації I , яку містять всі 24 метрики складності та якості разом, згідно формули для кількості інформації, наведеної вище, дорівнює 1. Вважатимемо також, що показник незнання складності та якості ПЗ зменшується в рівному ступені після опрацювання кожної метрики, тому кількість інформації кожної метрики якості $l_4 = I_m = \frac{1}{24} = 0,042$.

Представимо метрики із значеннями 4-х показників ефекту та ефективності у таблиці 2.

Таблиця 2. Ефективність та інформативність метрик програмного забезпечення на етапі проектування

№	Метрика	l_1	l_2	l_3	l_4	Ефективність A
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	Метрика Чепіна	1	0,333	0,333	0,042	0,107
2	Метрика Джилба (абсолютна)	1	0,333	0,333	0,042	0,107
3	Метрика Джилба (відносна)	1	0,333	0,333	0,042	0,107
4	Метрика Мак-Клура	1	1	1	0,042	0,190
5	Метрика Кафура	1	0,333	0,333	0,042	0,107
6	Очікувана LOC-оцінка	1	0,1665	0,1665	0,042	0,086
7	Метрика Холстеда	1	0,1665	0,1665	0,042	0,086
8	Метрика Маккейба	1	1	1	0,042	0,190
9	Прогнозована кількість операторів програми	1	0,1665	0,1665	0,042	0,086
10	Прогнозована оцінка складності інтерфейсів	1	0,4995	0,4995	0,042	0,128
11	Метрика зв'язності	1	0,666	0,666	0,042	0,148
12	Метрика зчеплення	1	1	1	0,042	0,190
13	Метрика звертання до глобальних змінни	1	0,4995	0,4995	0,042	0,128

(1)	(2)	(3)	(4)	(5)	(6)	(7)
14	Час модифікації моделей	1	0,333	0,333	0,042	0,107
15	Загальний час розроблення ПЗ	1	0,333	0,333	0,042	0,107
16	Кількість виявлених помилок при інспектуванні моделей	1	0,333	0,333	0,042	0,107
17	Час етапу проектування	1	0,333	0,333	0,042	0,107
18	Продуктивність розроблення ПЗ	1	0,333	0,333	0,042	0,107
19	Вартість перевірки якості	1	0,333	0,333	0,042	0,107
20	Вартість розроблення ПЗ	1	0,333	0,333	0,042	0,107
21	Вартість реалізації коду	1	0,333	0,333	0,042	0,107
22	Функційний розмір	1	1	1	0,042	0,190
23	Оцінка трудовитрат за моделлю Боєма	1	0,666	0,666	0,042	0,148
24	Оцінка тривалості проекту за Боємом	1	0,666	0,666	0,042	0,148

Діапазони значень метрик [3] мають дуже великий розкид — одне значення від іншого може відрізнятись в мільйон разів. Тому часто при опрацюванні значень метрик відбувається неврахування метрик з малими діапазонами значень. Оцінимо кількість інформації, яка втрачається при неврахуванні метрик з малими діапазонами значень (метрики Джилба (відносної), прогнозованої оцінки складності інтерфейсів, зв'язності, зчеплення, звертання до глобальних змінних, часу модифікації моделей, продуктивності розроблення ПЗ). Якщо кількість інформації кожної метрики якості $I_m = 0,042$, то кількість інформації семи метрик: $I_{grp} = 7 \cdot 0,042 = 0,294$.

Висновки. Отже, порівняно високу ефективність (значущість, інформативність) на етапі проектування програмного забезпечення мають метрики: Мак-Клура, Маккейба, зчеплення, функційного

розміру, зв'язності, оцінки трудовитрат та тривалості проекту за моделлю Боєма; низьку ефективність мають метрики: очікуваної LOC-оцінки, Холстеда, прогнозованої кількості операторів програми.

Очевидно, що неможливо не враховувати метрики з малими діапазонами значень, оскільки ефективність багатьох з них достатньо висока, тобто їх неврахування призведе до істотного зниження ефективності оцінювання складності та якості проекту і ПЗ. При неврахуванні впливу на оцінку складності та якості проекту і ПЗ метрик з малими діапазонами значень втрачається 29,4% значущої інформації.

Використані джерела

1. Орлов С.А. Технологии разработки программного обеспечения. Разработка сложных программных систем: Учебник для ВУЗов - СПб.: Питер, 2004. - 527 с.

2. Поморова О.В., Говорущенко Т.О., Тарасек С.Я. Аналіз та опрацювання метрик якості програмного забезпечення на етапі проектування // Вісник Хмельницького національного університету – Хмельницький: ХНУ, 2010. - № 1, с.54-63

3. Поморова О.В., Говорущенко Т.О., Онищук О.С. Оцінювання результатів проектування та прогнозування характеристик якості програмного забезпечення // Вісник Хмельницького національного університету - Хмельницький: ХНУ, 2011 - №2, с.168-178

4. И.Нежданов. Критерии оценки информации (важность, точность, значимость) // <http://www.police-russia.ru/showthread.php?t=44683>

5. Белик Т.В. Принципы формирования критериев эффективности // <http://lego.biuss.ru/paragraf/742>