

УДК 004.412:519.237.5

ВИЗНАЧЕННЯ СКЛАДНОСТІ ОБ'ЄКТНО-ОРІЄНТОВАНОГО ПРОЄКТУВАННЯ ЗА РАХУНОК ІДЕНТИФІКАЦІЇ КЛАСІВ ВЕБ ЗАСТОСУНКІВ РОЗРОБЛЕНИХ НА ОСНОВІ РНР ФРЕЙМВОРКІВ

Приходько А.С., аспірант, whiterandrek@gmail.com, ОНУ імені І.І.Мечникова

В наш час проблема визначення складності об'єктно-орієнтованого проєктування (ООП) програмного забезпечення (ПЗ), включаючи веб застосунки, які створюють за допомогою РНР фреймворків, є важливою, оскільки, по-перше, компанії прагнуть випускати якісне ПЗ. А, як відомо, на якість ПЗ впливає і складність його проєктування. По-друге, зараз при створенні веб застосунків активно використовують РНР фреймворки, які дозволяють робити розробку ПЗ швидшою.

Як відомо, проєктування в цілому і ООП зокрема є одним із важливих етапів розробки ПЗ. Буч окреслив чотири основні етапи процесу ООП [1]. Першим етапом в ООП є ідентифікація класів. На цьому етапі ключові абстракції в проблемному просторі ідентифікуються та позначаються як потенційні класи та об'єкти. І від того, наскільки вдалим буде ідентифікація класів, залежатиме подальша складність ООП і загальна якість ПЗ.

У [2] Чідамбер і Кемерер запропонували набір метрик ПЗ для ООП. Відповідно до [2], метрики Weighted Methods per Class (WMC), Depth of Inheritance Tree (DIT) та Number of Children (NOC) відноситься до першого етапу ООП, оскільки WMC є аспектом складності класу, і як DIT, так і NOC безпосередньо пов'язані з ієрархією класів. Тому метрики WMC, DIT та NOC можуть бути використані для оцінювання складності ООП за рахунок ідентифікації класів ПЗ. Але, як правило, ці метрики аналізуються окремо без врахування кореляції між ними, а прийнятні їх значення визначаються як на рівні класу, так і на рівні застосунку. Але, як відомо [3], існує достатньо значна кореляція між програмними метриками. Тому виникає потреба у визначенні складності ООП за рахунок ідентифікації класів, у тому числі веб застосунків, що створюються за допомогою РНР фреймворків, із врахуванням кореляцій між зазначеними метриками. А це можна робити за допомогою відповідних математичних моделей, які потребують побудови.

У [3] для оцінювання складності ООП через зв'язки між класами запропонована математична модель, яка побудована у вигляді еліпсу прогнозування в залежності від нормалізованих метрик Response for Class (RFC) та Coupling Between Objects (CBO). Беручи за основу запропонований в [3] підхід щодо оцінювання складності ООП, у якості математичної моделі для оцінювання складності ООП за рахунок ідентифікації класів веб застосунків, що розроблюються за допомогою РНР фреймворків, пропонується використовувати

рівняння еліпсоїду прогнозування в залежності від нормалізованих метрик WMC, DIT та NOC.

Побудову рівнянь зазначених еліпсоїдів прогнозування було здійснено за даними метрик WMC, DIT та NOC для 121 веб застосунку із відкритим кодом, що були створені із використанням наступних відомих PHP фреймворків: CakePHP, CodeIgniter, Laravel, Symfony та Yii. Значення метрики цих застосунків були отримані завдяки інструменту PhpMetrics (<https://phpmetrics.org/>). Розподіл отриманих тривимірних даних був перевірений на нормальність за допомогою тесту Мардіа (Mardia's test), який базується на багатовимірних ексцесі та асиметрії. Тест показав, що розподіл тривимірних даних метрик WMC, DIT та NOC не є гаусівським. А це вимагає будувати еліпсоїди прогнозування саме для нормалізованих метрик WMC, DIT та NOC. Нормалізацію даних було здійснено за тривимірним перетворенням Бокса-Кокса з компонентами

$$Z_j = x(\lambda_j) = \begin{cases} (X_j^{\lambda_j} - 1)/\lambda_j, & \text{якщо } \lambda_j \neq 0; \\ \ln(X_j), & \text{якщо } \lambda_j = 0. \end{cases} \quad (1)$$

В (1) Z_j – це гаусова змінна; λ_j є параметром перетворення Бокса-Кокса, $j = 1, 2, 3$. Ми позначили метрики WMC, DIT та NOC як X_1 , X_2 та X_3 , відповідно. Оцінки параметрів перетворення Бокса-Кокса були знайдені за методом максимальної правдоподібності: $\hat{\lambda}_1 = -0,177882$, $\hat{\lambda}_2 = -1,272192$ та $\hat{\lambda}_3 = 1,811640$.

Рівняння еліпсоїду прогнозування для нормалізованих метрик WMC, DIT та NOC має вигляд

$$(\mathbf{Z} - \bar{\mathbf{Z}})^T \mathbf{S}_Z^{-1} (\mathbf{Z} - \bar{\mathbf{Z}}) = \chi_{3,\alpha}^2, \quad (2)$$

де \mathbf{Z} є гаусовим випадковим вектором, $\mathbf{Z} = \{Z_1, Z_2, Z_3\}^T$; $\bar{\mathbf{Z}}$ – це вектор вибірових середніх, $\bar{\mathbf{Z}} = \{\bar{Z}_1, \bar{Z}_2, \bar{Z}_3\}^T$; N – це кількість точок даних; $\chi_{3,\alpha}^2$ є квантилем розподілу хі-квадрат із 3 ступенями свободи та рівнем значущості α ; \mathbf{S}_Z – це вибіркова коваріаційна матриця

$$\mathbf{S}_Z = \begin{pmatrix} S_{Z_1Z_1} & S_{Z_1Z_2} & S_{Z_1Z_3} \\ S_{Z_2Z_1} & S_{Z_2Z_2} & S_{Z_2Z_3} \\ S_{Z_3Z_1} & S_{Z_3Z_2} & S_{Z_3Z_3} \end{pmatrix}, \quad (3)$$

$$\text{де } S_{Z_qZ_r} = \sum_{i=1}^N [Z_{q_i} - \bar{Z}_q][Z_{r_i} - \bar{Z}_r], \quad q, r = 1, 2, 3.$$

Ліва частина в (2) – це квадрат відстані Махаланобісу (КВМ) для нормалізованих метрик WMC, DIT та NOC. У нашому випадку при обчисленні значення КВМ потрібно використовувати вектор \bar{Z} з компонентами які відповідно дорівнюють $-0,17788$, $-1,27219$ та $1,81164$, а також наступну зворотню матрицю до коваріаційної матриці (3):

$$S_Z^{-1} = \begin{pmatrix} 3.877 & 8.398 & -9.411 \\ 8.398 & 139.516 & -111.477 \\ -9.411 & -111.477 & 194.605 \end{pmatrix}. \quad (4)$$

Для визначення складності ООП за рахунок ідентифікації класів веб застосунків розроблених на основі РНР фреймворків пропонується використовувати два рівняння (2) з двома значеннями квантилю $\chi_{3,\alpha}^2$ для рівня значущості 0,05 та 0,005 – це 7,82 та 12,84, відповідно. Тоді ми можемо визначати складність ООП за рахунок ідентифікації класів для певного застосунку за (2) у разі, якщо КВМ його нормалізованих метрик WMC, DIT та NOC буде менше або дорівнювати 12,84. Якщо КВМ менше або дорівнює 7,82, то складність ООП за рахунок ідентифікації класів не висока. У разі, якщо КВМ більше 7,82 та менше або дорівнює 12,84, то складність ООП за рахунок ідентифікації класів висока. Застосування (2) з (1) та (4) до даних метрик WMC, DIT і NOC з 121 веб застосунку розроблених на основі РНР фреймворків показало, що складність ООП за рахунок ідентифікації класів цих застосунків не висока.

Висновок. Удосконалено рівняння еліпсоїду прогнозування для визначення складності ООП за рахунок ідентифікації класів веб застосунків, що розробляються на основі РНР фреймворків, на основі нормалізованих метрик WMC, DIT та NOC. Ці рівняння дозволяють врахувати кореляцію між зазначеними метриками при визначенні складності ООП за рахунок ідентифікації класів. В подальшому планується пошук нових даних для підтвердження працездатності запропонованих рівнянь.

Список використаних джерел

1. Booch G. Object oriented design with applications. Redwood City, CA: Benjamin/Cummings, 1991. 580 p.
2. Chidamber S, Kemerer C. A metrics suite for object oriented design. IEEE Transactions on Software Engineering. 1994; 20(6): 476-493. doi: 10.1145/118014.117970
3. Prykhodko S, Prykhodko N, Smykodub T. A joint statistical estimation of the RFC and CBO metrics for open-source applications developed in Java. 2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT); 10-12 November 2022; Lviv, Ukraine; IEEE; 2022. p. 442-445. doi: 10.1109/CSIT56902.2022.10000457