

РАННЄ ОЦІНЮВАННЯ РОЗМІРУ CMS МОВОЮ PHP

Латанська Л.О., к.ф.-м.н., доц.¹; Савчук Г.Ю.²

^{1,2} Національний університет кораблебудування імені адмірала Макарова

^{1,2} Україна, Миколаїв

¹ llatanskaya@gmail.com; ² greg.savchuk12@gmail.com

Анотація. В роботі було побудовано нелінійну регресійну модель для оцінювання кількості строк коду CMS мовою PHP.

Ключові слова: LOC оцінювання; веб-застосунок; нелінійна регресійна модель; розробка програмного забезпечення.

Вступна частина. Визначення трудомісткості розробки на початкових етапах роботи над проектом є важливою задачею проектних менеджерів. Існують різні підходи для визначення трудомісткості, зокрема використання математичних моделей таких як, наприклад, СОСОМО II. Застосування моделі СОСОМО II передбачає знання про розмір програмного забезпечення у строках коду, який в більшості випадків невідомий. Саме тому існує потреба в ранньому оцінюванні розміру програмного забезпечення.

CMS (Content Management System) – це системи управління контентом, які дозволяють створювати, редагувати, організовувати та керувати цифровим вмістом на сайті без глибоких знань програмування. CMS є програмними платформами, що надають користувачам зручний інтерфейс і безліч вбудованих функцій для налаштування та адміністрування сайту. За допомогою CMS можна легко додавати нові сторінки, оновлювати інформацію, керувати дизайном, структурою та функціональністю веб-сайту. Історично склалось, що мова PHP є основною мовою веб-розробки й більша частина CMS розроблені саме цією мовою.

На сьогодні відомі як лінійні, так і нелінійні регресійні моделі для визначення розміру програмного забезпечення для таких мов програмування як C++, Java, Visual Basic, PHP [1-6]. Вищезазначені моделі будувались з використанням метрик діаграми класів (у різних комбінаціях), серед яких: загальна кількість класів, загальна кількість зв'язків, середня кількість атрибутів на клас, сума середньої кількості класів, на які впливає даний клас, і середньої кількості класів, з яких даний клас отримує ефекти, середня кількість методів, глибина дерева наслідування. У роботі [6] автори досліджували PHP фреймворки та визначили, що для отримання достовірних оцінок розміру існує необхідність будувати нелінійні регресійні моделі під кожний окремий фреймворк. Ми схильні вважати, що це також справедливо й для проектів різних типів, зокрема для CMS, що й слугує підставою для проведення досліджень.

Метою роботи є побудова нелінійної регресійної моделі для оцінювання розміру CMS мовою PHP у тисячах строк коду з використанням нормалізуючого перетворення десятикового логарифму.

Основна частина. Побудова нелінійної регресійної моделі на першому етапі вимагає пошуку емпіричних даних. Серед розміщених на сайті GitHub (<https://github.com/>) було знайдено 45 проектів CMS мовою PHP. Для отримання метрик програмного забезпечення було використано інструмент PhpMetrics. Для побудови нелінійної регресійної моделі використовувались метрики діаграми класів такі, як розмір веб-застосунку Y у тисячах строк коду, кількість класів X_1 , середня кількість методів у класі X_2 , глибина дерева наслідування X_3 . Наведений перелік метрик було обрано через ряд причин, серед яких: можливість

отримання цих метрик на ранній стадії роботи над проектом з діаграми класів, відсутність мультиколінеарності між метриками. Для визначення викидів у даних використовували квадрат відстані Махаланобіса [7].

Отримана нелінійна регресійна модель для CMS мовою PHP має вигляд: $Y = 10^{\varepsilon + \hat{b}_0} X_1^{\hat{b}_1} X_2^{\hat{b}_2} X_3^{\hat{b}_3}$, де ε – гаусівська випадкова величина. Оцінки параметрів $\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3$ дорівнюють -1,9927, 0,9663, 1,5259 та -0,4627 відповідно. Якість побудованої моделі було оцінено за допомогою розрахунку множинного коефіцієнту детермінації R^2 , середньої величини відносної похибки $MMRE$ та відсотку прогнозованих результатів $PRED(0,25)$. Значення R^2 , $MMRE$ і $PRED(0,25)$ для створеної моделі дорівнюють відповідно 0,970, 0,148 і 0,806, що свідчить про прийнятну якість отриманої моделі [7].

Висновки. В роботі було побудовано нелінійну регресійну модель, для оцінювання розміру CMS мовою PHP, яка залежить від трьох метрик діаграми класів. Отримана модель має задовільні оцінки якості. В подальшому планується застосувати більш потужні нормалізуючі перетворення для отримання більш якісної моделі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Kiewkanya, M., & Surak, S. (2016). Constructing C++ software size estimation model from class diagram. *2016 IEEE 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 1-6). Khon Kaen, Thailand. <https://doi.org/10.1109/jcsse.2016.7748880>
- [2] Prykhodko, N. V., & Prykhodko, S. B. (2018). The non-linear regression model to estimate the software size of open source Java-based systems. *Radio Electronics, Computer Science, Control*, 3, 158-166. <https://doi.org/10.15588/1607-3274-2018-3-17>
- [3] Prykhodko, N. V., & Prykhodko, S. B. (2018). Non-linear regression model to estimate the software size of Vb-based information systems. *Information Technology and Computer Engineering*, 43(3), 37-42. <https://doi.org/10.31649/1999-9941-2018-43-3-37-42>
- [4] Tan, H.B.K., Zhao, Y., & Zhang, H. (2006). Estimating LOC for information systems from their conceptual data models. *Software Engineering : 2006 28th International Conference (ICSE '06)* (pp. 321-330). Shanghai, China. <https://doi.org/10.1145/1134285.1134331>
- [5] Приходько, С.Б., Приходько, Н.В., Фаріонова, Т.А., & Ворона, М.В. (2020). Трьохфакторна нелінійна регресійна модель для оцінювання розміру PHP-застосунків з відкритим кодом. *Науковий журнал «Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Технічні науки»*, 31 (70), 124-131. doi: 10.32838/2663-5941/2020.1- 1/23
- [6] Prykhodko, S., Shutko, I., & Prykhodko, A. (2022). Early size estimation of web apps created using codeigniter framework by nonlinear regression models. *Radioelectronic and computer systems*, 3, 84-94. <https://doi.org/10.32620/reks.2022.3.06>
- [7] Prykhodko, N.V., & Prykhodko, S.B. (2018). Constructing the non-linear regression models on the basis of multivariate normalizing transformations. *Electronic modeling*, 6 (40), 101-110. doi: 10.15407/emodel.40.06.101

Latanska Liudmyla, Savchuk Hryhorii

Early LOC estimation of PHP-based CMS

Abstract. In work a nonlinear regression model for estimating LOC of PHP-based CMS was created.

Keywords: LOC estimation; web application; nonlinear regression model; software development.