

УДК 004.413:338.5

В.А. Хоменко

МЕТОД ОЦІНЮВАННЯ ПОВТОРНО ВИКОРИСТОВУВАНИХ КОМПОНЕНТІВ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

The study proposes the evaluation method of reusable software components. The method provides the automated acquisition of properties values based on the directly measured metrics. To fulfill this method, we propose to develop two-level "property—metrics" models that support calculation of properties values based on the direct metrics values. By analyzing aspects of software components based on the 3C-model "conception—construction—context", the ontology of appropriated properties is developed. The specific technique is proposed to create and to investigate the model properties. This technique encompasses collecting the empirical data of evaluations, checking model dependencies using statistical analysis and modeling the properties calculation. The proposed method enables reducing the efforts for the software evaluation aimed at selecting reusable software components at the non-systematical reuse processes.

Вступ

Повторне використання залишається найбільш перспективним способом зниження витрат і підвищення продуктивності розроблення програмного забезпечення (ПЗ) [1, 2]. Повторне використання застосовується більшістю виробників ПЗ і дає можливість скоротити час виходу продуктів на ринок у 2–5 разів, знизити кількість помилок і вартість супроводження в 5–10 разів, а загальну вартість розробки – від 15 до 75 % [2, 3]. Вказані значення забезпечуються за рахунок скорочення витрат на пошук розв'язків задач, які виникають при розробленні ПЗ, мінімізації витрат на кодування, підвищення надійності роботи ПЗ, оскільки використовуються готові, добре тестовані компоненти. Важливість повторного використання при розробленні ПЗ знайшла також відображення у тому, що створено групу моделей життєвого циклу ПЗ, яку засновано на повторному використанні.

При застосуванні повторного використання ПЗ необхідно створювати повторно використовувані компоненти (ПВК). За способом реалізації процесів виробництва ПВК повторне використання може бути систематичним і несистематичним. Сутність систематичного повторного використання полягає в створенні "з нуля" ПВК для визначеного домену та їх подальшому багаторазовому використанні при розробленні ПЗ. Несистематичне повторне використання – це використання компонентів вже існуючого (успадкованого) ПЗ у розробленні нового ПЗ. Цей тип повторного використання припускає використання елементів раніше розробленого ПЗ через їх виділення з наслідуваного ПЗ, переробку і/або налаштування та інтеграцію в нове ПЗ.

Постановка задачі

Під час несистематичного повторного використання ПВК створюються на основі успадкованого ПЗ, а доцільність та ефективність створення ПВК визначаються відповідними витратами. Оцінювання властивостей компонента ПЗ – кандидата в ПВК потребує великих обсягів робочого часу висококваліфікованих інженерів і виконується інтуїтивно через розгляд та аналіз компонентів успадкованого ПЗ, в першу чергу – кодів програм. Мета роботи – запропонувати метод автоматизації процесів оцінювання компонентів успадкованого ПЗ з точки зору повторного використання.

Вихідні положення

Створення нового ПЗ із застосуванням повторного використання складається з двох процесів (рис. 1). Сутність першого полягає у розробленні ПЗ, а другого – у виробництві ПВК.

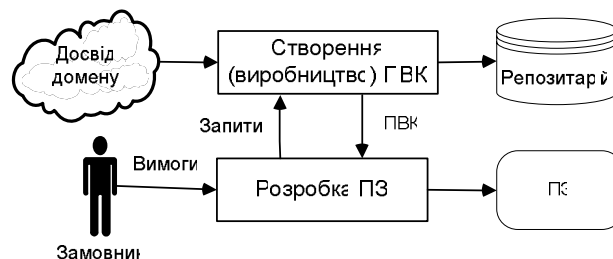


Рис. 1. Розроблення нового ПЗ із застосуванням повторно використовуваних компонентів

При несистематичному повторному використанні в процесі відбору компонентів ПЗ для ПВК програміст здійснює аналіз компонентів успадкованого ПЗ та ухвалює рішення про їх придатність до повторного використання (рис. 2).

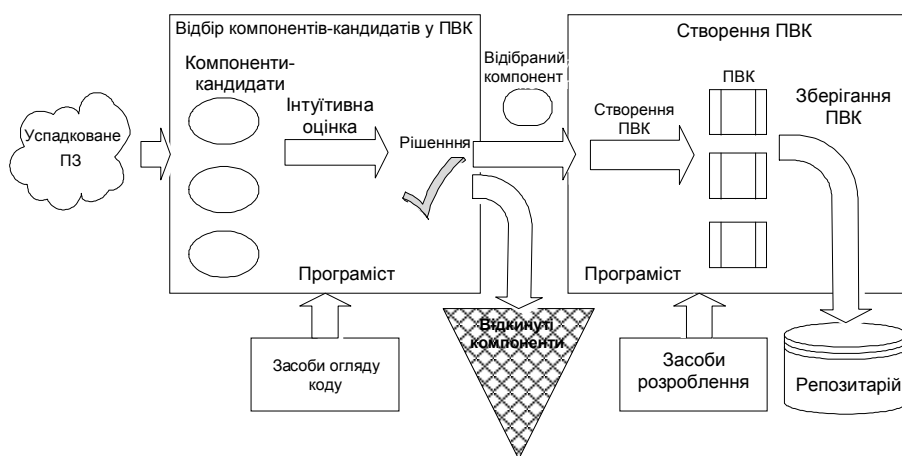


Рис. 2. Відбір компонентів ПЗ – кандидатів у ПВК

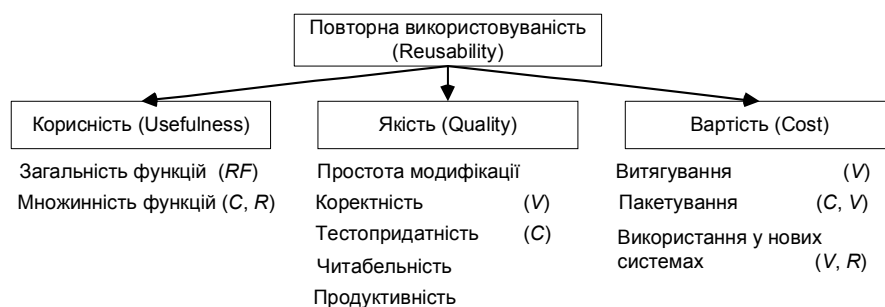


Рис. 3. Модель факторів повторного використання

Відбір компонентів зазвичай будується на припущенні, що створення ПВК не повинно вимагати витрат, більших за необхідні для створення аналогічного компонента “з нуля” [4]. Витрати можна оцінювати прямо, використовуючи відомості про вартість необхідних процесів повторного використання, або непрямо, використовуючи знання про те, наскільки властивості компонента ПЗ відповідають функціональним і нефункціональним вимогам до ПВК. Тому для відбору компонентів ПЗ – кандидатів у ПВК нині використовуються два типи моделей:

вартості (рішення про повторне використання ухвалюється на основі оцінки вартості виконання процесів, пов’язаних із перетворенням компонентів ПЗ на ПВК [5, 6]);

властивостей (рішення про повторне використання ухвалюється на основі оцінок властивостей, які відображають придатність компонента до повторного використання [5]).

Модель вартості будується через ідентифікацію дій (операцій), які відповідають кожному типу повторного використання (розроблення “з нуля”, несистематичне, систематичне повторне використання), і асоційованої з ними

вартості (витрат), необхідної для виконання цих дій. Далі, розподіляючи загальну вартість розробки, засновану на повторному використанні, на три частини (створення продукту, конструювання базового набору ПВК, підтримка інфраструктури) і будуючи для процесів повторного використання сценарії у вигляді схем переходів у двовимірній моделі “перетворення–перенесення”, вартість відповідних процесів визначають як суму оцінок вартостей, що входять до сценарію базових операцій. Для несистематичного повторного використання вартість процесів має вигляд

$$C_{CA}(0, S_i) + C_{WB}(S_i, T_i),$$

де C_{CA} – вартість операцій видобування компонента S_i із успадкованого ПЗ; C_{WB} – вартість модифікації компонента ПЗ S_i (white-box reuse) у ПВК T_i .

При використанні моделей вартості, з урахуванням великого різноманіття як вихідних компонентів ПЗ (S_i), так і операцій, які потрібно виконати, похибка оцінки витрат може бути дуже значною.

Модель властивостей створюється на основі PVB-підходу (property-values based) асоціюванням з вимогами до компонентів так званих складних атрибутів [5] і деталізацією їх за допомогою простих атрибутів, що вимірюються. У праці [7] PVB-підхід використовується для відбору компонентів ПЗ – кандидатів до ПВК при несистематичному використанні. Модель властивостей компонентів містить:

- модель факторів повторного використання (рис. 3), яка складається з трьох факторів (корисності, якості, вартості ПВК) і субфакторів;

- чотири прямі метрики – цикломатичну складність (C), регулярність (R), обсяг (V), час-

тоту використання (*RF*) [8], які поставлені у відповідність до деяких із субфакторів.

Будуючи модель, автори припускають існування залежностей між впровадженими факторами й вибраними метриками, проте задача формалізації та підтвердження цих залежностей не розв'язується.

Отже, аналіз моделей, які застосовуються для відбору компонентів ПЗ – кандидатів до ПВК, показує, що для несистематичного повторного використання доцільно застосовувати комбінований підхід, основу якого становить модель властивостей, а при формуванні множини властивостей використовується модель вартості.

Сутність методу

За основу методу пропонується взяти PVB-підхід, який дає можливість під час відбору компонентів успадкованого ПЗ подавати їх через властивості, які відображають потенційну придатність цих компонентів до повторного використання. Реалізація методу експертного оцінювання компонентів ПЗ потребує виконання двох процесів:

- 1) оцінювання властивостей, суттєвих в аспекті повторного використання;
- 2) ухвалення рішення про створення ПВК або відкидання компонента.

З урахуванням трудомісткості першого процесу та можливості виконання вимірювань у ПЗ [9] пропонується отримувати оцінки всіх або деяких властивостей автоматично і відтак скоротити обсяг оцінювання властивостей компонента. Схема роботи експерта під час такого оцінювання показана на рис. 4.

Оскільки обчислювання значень властивостей компонентів, і відповідно, автоматизація експертної оцінки, має здійснюватися на основі моделей властивостей, то необхідно забезпечити формалізацію і адекватність цих моделей.

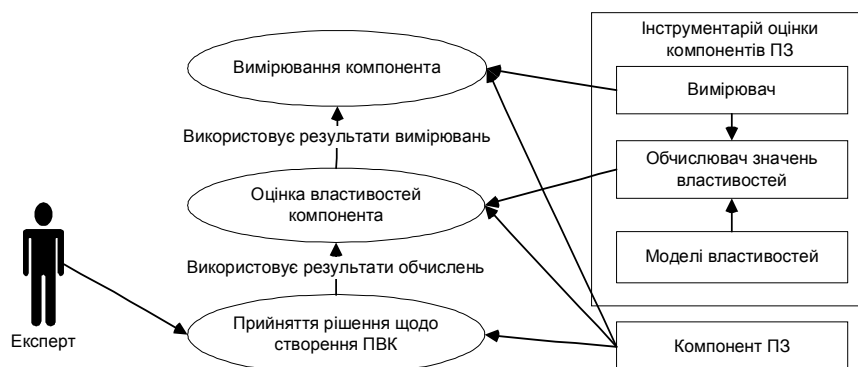


Рис. 4. Процеси автоматизованого експертного оцінювання

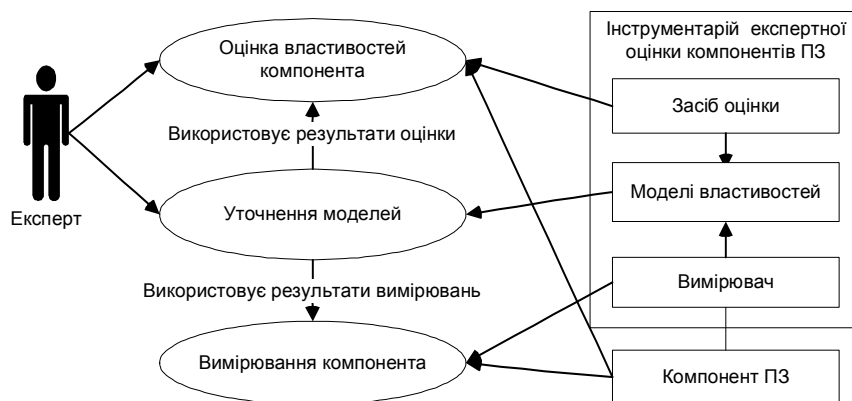


Рис. 5. Процеси та засоби уточнення моделей властивостей

Пропонується забезпечувати формалізацію і адекватність таких моделей їх уточненням за допомогою досвідчених експертів, які володіють знаннями щодо повторного використання. Це потребує послідовного виконання трьох процесів: оцінки властивостей компонентів, вимірювання компонентів та уточнення моделей (рис. 5).

Уточнення моделей властивостей має виконуватися після проведення достатньої кількості оцінок і вимірювань на основі значень, які зберігаються у базі даних.

Властивості компонентів ПЗ, суттєві з точки зору повторного використання

Для побудови моделей властивостей, суттєвих з точки зору повторного використання, пропонується використовувати 3С-модель компонентів ПЗ [10]. Сутність 3С-моделі компонента ПЗ полягає в тому, що онтологія властивостей компонентів ПЗ будується з урахуванням трьох аспектів: концепції (Conception), конструкції (Construction) і контексту (Context).

Концепція програмного компонента відображає його функціональність. Доменний експерт сприймає концепцію у вигляді функцій, які моделюють визначені об'єкти і процеси предметної області. З точки зору повторного використання функціональність повинна бути корисною та загальною. Через це в аспекті концепції компонента суттєвими є властивості корисності та загальності. Корисність відображає ступінь задоволення вимог, які користувачі висувають до компонента. Загальність відображає широту можливого застосування ПВК.

Контекст відображає вимоги до середовища використання ПВК у складі ПЗ. У ЗС-моделі можна виділити контекст концепції (семантичний контекст) і контекст конструкції (синтаксичний контекст). При створенні ПВК семантичний контекст формується програмістом як логіка його взаємодії з іншими компонентами у прикладному домені. Синтаксичний контекст формується як інтерфейс ПВК. Тому в аспекті контексту суттєвими є властивості семантичної та синтаксичної незалежності, повнота інтерфейсу. Семантична незалежність відображає здатність компонента давати при виконанні фіксований результат незалежно від місця, де він оголошується або викликається. Синтаксична незалежність відображає здатність компонента використовуватися без змінення контексту, в який він поміщається. Повнота інтерфейсу відображає те, що інтерфейс компонента надає всі відкриті функції компонента.

Конструкція програмного компонента відображає одну з реалізацій концепції. Отже, для конструкції суттєвими в аспекті повторного використання є такі властивості:

- достатність — повнота реалізації концепції;
- примітивність — включення в конструкцію базових операцій концепції, які дають можливість використовувати компонент стандартним чином;

- модульність — модульна внутрішня структура компонента, яка спрощує його розуміння та модифікацію;

- структурованість — використання в компоненті структурних операторів та їх складність, яка впливає на його розуміння та модифікацію;

- слідування — можливість мати в процесі розробки “слід” від однієї фази життєвого циклу до іншої та використовувати ефект інклюзії;

- надійність — стійкість поведінки компонента;

- ефективність — ефективне виконання компонента в рамках всього спектра застосувань, які передбачені для нього специфікацією вимог;

- переносимість — забезпечення працездатності компонента в різних апаратних і програмних середовищах;

- придатність до класифікації — здатність ефективного розміщення та пошуку компонента відповідно до заданої класифікації;

- компактність — розмір зберігання компонента;

- зрозумілість — самодокументованість компонента;

- простота модифікації — легкість внесення змін у конструкцію компонента;

- коректність — відсутність синтаксичних і семантичних помилок у тексті компонента;

- тестопридатність — можливість ефективного тестування компонента.

Методика побудови і дослідження моделей властивостей ПВК

Для реалізації запропонованого методу необхідно побудувати дворівневі моделі властивостей, суттєвих з точки зору повторного використання [11]: перший рівень — властивостей, що оцінюються, другий — метрик, що вимірю-

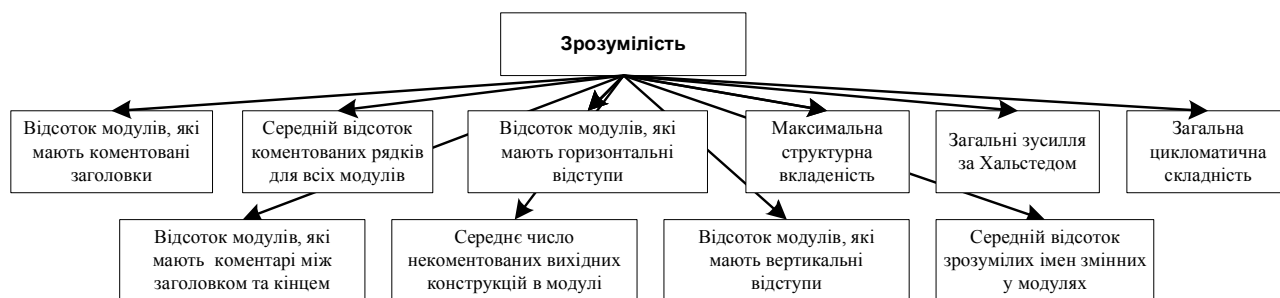


Рис. 6. Модель властивості “Зрозумілість”

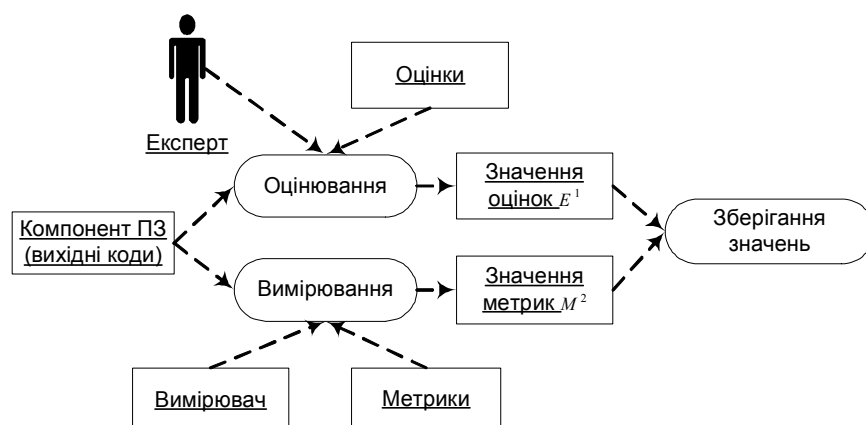


Рис. 7. Схема реалізації етапу накопичення даних

ються. Для цього можна використовувати метод GQM (“ціль–питання–метрика”) [12]. Приклад моделі властивості “Зрозумілість” поданий на рис. 6.

Методика, яка пропонується для дослідження адекватності моделей властивостей та уточнення залежностей між властивостями компонентів і метриками, складається з трьох етапів:

- 1) накопичення даних у вигляді результатів експертного оцінювання та вимірювання компонентів ПЗ;
- 2) статистичного аналізу отриманих даних для уточнення залежностей між оцінками та метриками;
- 3) моделювання роботи обчислювача значень властивостей.

Сутність першого етапу полягає в отриманні значень оцінок властивостей і метрик компонентів ПЗ за допомогою їх оцінки досвідченими експертами та вимірювання початкових кодів компонентів. Для реалізації експертного оцінювання та вимірювання компонентів ПЗ застосовуються засоби, які належать до класу систем Computer Aided Empirical Software Engineering (CAESE). Початкові коди кожного компонента ПЗ надходять експерту, який формує значення експертних оцінок $\{e_1, e_2, \dots, e_m\} \in E^1$ методом безпосередньої оцінки, та на вимірювач, який визначає значення прямих метрик $\{m_1, m_2, \dots, m_n\} \in M^2$. Отримані значення зберігаються в базі даних разом із додатковою інформацією, потрібною для подальшої обробки результатів (рис. 7).

Для реалізації експертного оцінювання формується набір оцінок, який відповідає множині властивостей, суттєвих з точки зору повторного використання. Для кожної оцінки ви-

значаються назва, семантика та шкала оцінювання у вигляді множини бальних значень. Під час оцінювання конкретного компонента множина значень кожної оцінки подається експерту у вигляді впорядкованого списку значень, з якого він вибирає те значення, яке, на його думку, характеризує відповідну властивість компонента. Наприклад, для оцінки властивості “Зрозумілість” експерт може вибирати значення з переліку:

“незадовільна”, “задовільна”, “добра”, “відмінна”.

Для реалізації вимірювань компонентів ПЗ використовується вимірювач початкових кодів програм, який забезпечує отримання значень прямих метрик.

Сутність другого етапу полягає в уточненні залежностей між оцінками і метриками через аналіз значень, отриманих на першому етапі, із застосуванням методів:

- первинного статистичного аналізу для визначення законів розподілення для кожної метрики та оцінки окремо;
- кореляційного аналізу для визначення залежностей між кожною метрикою та експертною оцінкою з подальшим відсіюванням пар, які не показують залежностей;
- регресійного аналізу пар метрика–експертна оцінка, відібраних у результаті аналітичного знаходження видів функції залежності та їх коефіцієнтів.

Результатом другого етапу є моделі властивостей компонентів ПЗ, уточнені з використанням емпіричних даних.

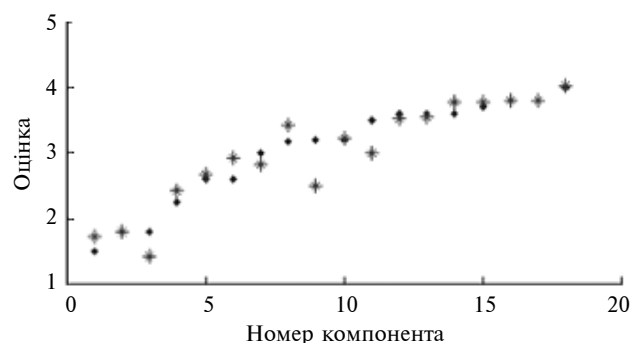


Рис. 8. Порівняльний графік експертних оцінок та обчислених значень властивості “Зрозумілість”: * – значення оцінок; ♦ – обчислені значення властивості

Сутність третього етапу полягає в моделюванні роботи обчислювача значень властивостей, яке може виконуватися різними засобами. Приклад результатів використання адаптивної мережі нечіткого виводу математичного пакета MatLab наведений на рис. 8.

Висновки

Метод, що пропонується у статті, дає можливість знизити витрати на оцінювання компонентів ПЗ при їх відборі для створення ПВК за рахунок автоматизації отримання значень властивостей, суттєвих з точки зору повторного використання. Автоматичне обчислювання значень властивостей засновано на використанні дворівневих моделей “властивість—прямо вимірювані метрики” і забезпечується спеціальними засобами.

Для реалізації методу запропоновано онтологію властивостей компонентів ПЗ, суттєвих з точки зору повторного використання, побудовано із застосуванням 3С-моделі з урахуванням впливу властивостей на витрати для виконання процесів повторного використання. Показано побудову дворівневих моделей із застосуванням відомого методу GQM.

Для перевірки адекватності розроблених моделей та їх уточнення розроблено й апробовано методику, яка включає статистичну обробку та аналіз емпіричних даних, що накопичуються при експертному оцінюванні, і моделювання роботи засобів обчислення значень властивостей.

Подальші дослідження стосовно запропонованого методу спрямовані на розвиток і уточнення онтології та моделей властивостей компонентів, суттєвих з точки зору повторного використання.

1. *V.R. Basili and B. Boehm*, “COTS – Based System Top 10 List”, *Software management*, May, 2001, pp. 124–137.
2. *I. Jacobson et al.*, *Software Reuse. Architecture, Process and Organization for Business Success*. Pearson Education Asia, 2002, 497 pp.
3. *M.L. Griss*, “Software reuse: From library to factory”, *IBM Syst. J.*, vol. 32, no. 4, pp. 548–565, 1993.
4. *T. Biggerstaff and C. Richter*, “Reusability Framework, Assessment, and Directions”, *IEEE Software*, March, pp. 42–57, 1987.
5. *B.W. Boehm*, “Improving Software Productivity”, *Computer*, vol. 20, no. 9, pp. 43–57, 1987.
6. *A. Tomer et al.*, “Evaluating Software Reuse Alternatives: A Model and Its Application to an Industrial Case Study”, *IEEE Trans. Soft. Eng.*, vol. 30, no. 9, pp. 111–118, 2004.
7. *G. Caldiera and R. Basili*, “Identifying and Qualifying Reusable Software Components”, *IEEE Comput.*, vol. 24, no. 2, pp. 61–70, 1991.
8. *N.E. Fenton and S.L. Pfleeger*, *Software Metrics: A Rigorous and Practical Approach*. Cambridge University Press, 1996, 638 pp.
9. *ISO/IEC 9126-1*, *Information Technology – Software Product Quality – Part 1: Quality Model*, International Organization for Standardization, 2001, 25 pp.
10. *R.K. Ray and H.M. Levy*, “A compositional model for Software Reuse”, *Comput. J.*, vol. 32, no. 4, pp. 312–322, 1989.
11. *Хоменко В.А.* Использование метрик для повышения эффективности повторного использования программного обеспечения // Проблемы программирования. – 2002. – № 1-2. – С. 131–135.
12. *V.R. Basili and H.D. Rombach*, “The TAME project: Toward improvement-oriented software environments”, *IEEE Trans. Soft. Eng.*, vol. 14, no. 6, pp. 758–73, 1988.