УДК 004.412:519.237
DOI https://doi.org/10.32782/2663-5941/2026.1.2/26

Макарова Л.М.
https://orcid.org/0000-0003-2903-3001
Національний університет кораблебудування імені адмірала Макарова

Пухалевич А.В.
https://orcid.org/0000-0002-8827-3251
Національний університет кораблебудування імені адмірала Макарова

Бризгалов М.В.
https://orcid.org/0009-0007-4728-3334
Національний університет кораблебудування імені адмірала Макарова

Камінський С.С.
https://orcid.org/0009-0006-1451-5779
ТОВ «Аджакс Сістемс Манюфекчурінг»

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ДЛЯ ОЦІНЮВАННЯ СКЛАДНОСТІ ОБ'ЄКТНО-ОРІЄНТОВАНОГО ПРОЄКТУВАННЯ JAVA-ЗАСТОСУНКІВ З ВИКОРИСТАННЯМ МЕТРИК RFC ТА CBO

У роботі розглянуто проблему оцінювання складності об'єктно-орієнтованого проєктування Java-застосунків на основі програмних метрик RFC (Response for a Class) та CBO (Coupling Between Objects). Актуальність даної проблеми зумовлена необхідністю своєчасного виявлення архітектурних аномалій і надмірної складності програмного забезпечення на ранніх етапах розробки, оскільки такі відхилення істотно впливають на якість, підтримуваність та вартість програмних продуктів. Існуючі підходи до аналізу складності проєктування часто не забезпечують достатньої достовірності результатів та не враховують статистичні особливості розподілу значень метрик, що може призводити до хибних висновків щодо стану архітектури застосунку.

З метою розв'язання цієї проблеми в роботі запропоновано математичну модель для оцінювання складності об'єктно-орієнтованого проєктування Java-застосунків, яка базується на побудові еліпса прогнозування у просторі метрик RFC та CBO. Для підвищення достовірності оцінювання застосовано багатовимірне перетворення Бокса–Кокса, що дозволяє нормалізувати статистичні дані, а також застосовано методи багатовимірного статистичного аналізу, зокрема використання відстані Махаланобіса для виявлення викидів. Запропонований підхід дає змогу більш точно ідентифікувати застосунки з аномальною архітектурною складністю та зменшити вплив шуму вхідних даних на результати аналізу.

На основі розробленої математичної моделі створено програмне забезпечення для автоматизованої оцінки складності Java-застосунків. Проведене тестування програмного забезпечення підтвердило ефективність запропонованого підходу та його переваги порівняно з традиційними методами аналізу метрик, зокрема у частині виявлення викидів та підвищення достовірності оцінювання складності об'єктно-орієнтованого проєктування.

Отримані результати можуть бути використані розробниками та менеджерами програмних проєктів для контролю архітектурної складності Java-застосунків, оптимізації процесів проєктування та підтримки програмного забезпечення, а також для зменшення ризиків, пов'язаних із зростанням складності та зниженням якості програмних продуктів.

Ключові слова: *складність проєктування, Java-застосунок, метрики RFC та CBO, еліпс прогнозування, перетворення Бокса–Кокса, відстань Махаланобіса, виявлення викидів.*

© Макарова Л.М., Пухалевич А.В., Бризгалов М.В. та ін., 2026
Стаття поширюється на умовах ліцензії відкритого доступу CC BY 4.0

Постановка проблеми. У сучасних умовах активного розвитку інформаційних технологій та широкого застосування об'єктно-орієнтованих підходів у програмуванні, зростання складності програмних застосунків стає суттєвою проблемою. Збільшення обсягу коду та кількості взаємозв'язків між компонентами зумовлює потребу в більш точних методах аналізу й прогнозування архітектурних характеристик програмного забезпечення (ПЗ) [1].

Ключову роль у цьому процесі відіграють метрики ПЗ, зокрема показники зв'язності та відповідальності класів, які найбільш повно характеризують структурну складність об'єктно-орієнтованих систем [2, 3]. Водночас, попри наявність різних підходів до оцінювання якості проєктування, сучасна інженерія ПЗ не має універсальної системи показників, здатної комплексно враховувати взаємозалежності між метриками. Це ускладнює точне прогнозування складності та своєчасне виявлення архітектурних аномалій.

Ситуацію додатково ускладнює розробка програмних продуктів у стислі терміни з орієнтацією переважно на бізнес-вимоги, коли вибір технологій часто базується на зовнішніх чинниках або попередньому досвіді команд. Такий підхід може призводити до надмірного ускладнення ПЗ, проблем із супроводом і зростання витрат на підтримку, особливо в розподілених та відкритих проєктах.

Існуючі моделі складності проєктування ПЗ здебільшого недостатньо враховують багатофакторні залежності між метриками. Тому актуальним є вдосконалення підходів до оцінювання складності об'єктно-орієнтованого проєктування, зокрема шляхом застосування моделей на основі побудови двох еліпсів прогнозування відповідно. Це дає змогу виявляти аномальні значення метрик і класифікувати проєкти за рівнем складності, підвищуючи достовірність оцінювання складності Java-застосунків.

Аналіз останніх досліджень і публікацій. Оцінювання складності об'єктно-орієнтованого проєктування ПЗ є важливим етапом у процесі аналізу якості архітектури, оскільки дозволяє своєчасно виявляти потенційні аномалії, що негативно впливають на підтримуваність, масштабованість та подальший розвиток програмних продуктів. Для розв'язання цього завдання у сучасних дослідженнях застосовуються різні математичні та статистичні моделі, побудовані на основі програмних метрик, зокрема метрик RFC та CBO, які є одними з ключових показників складності об'єктно-орієнтованого ПЗ.

Існуючі підходи до оцінювання складності архітектури ПЗ можна умовно поділити на декілька груп, які містять: лінійні та нелінійні регресійні моделі [4], статистичні моделі виявлення викидів [5], методи багатовимірного аналізу [6], а також моделі на основі машинного навчання [7]. Кожен із цих підходів має свої переваги та обмеження, які впливають на результати аналізу. Зокрема, лінійні регресійні моделі є простими в інтерпретації, однак часто не здатні адекватно описувати складні взаємозв'язки між метриками об'єктно-орієнтованого проєктування. Дослідження свідчать, що розподіли значень метрик RFC та CBO, а також залежності між ними, у більшості випадків мають нелінійний і асиметричний характер, що знижує ефективність класичних лінійних підходів.

У низці робіт запропоновано використовувати статистичні методи нормалізації та багатовимірного аналізу для підвищення достовірності оцінювання складності проєктування ПЗ. Зокрема, у [8] представлено удосконалений підхід до спільної статистичної оцінки метрик RFC та CBO для Java-застосунків з відкритим кодом. Запропонована модель базується на застосуванні перетворення Бокса–Кокса з метою нормалізації даних та побудові еліпсів прогнозування з використанням χ^2- та F-розподілів. Такий підхід дозволяє враховувати кореляцію між метриками та ефективно виявляти аномальні значення, що свідчать про надмірну складність окремих компонентів програмної системи.

Разом з тим, модель, запропонована в [8], не враховує відносне перетворення значень метрик із використанням кількості класів NCL, що може призводити до значної варіативності вхідних даних. В нашому випадку, значення метрик RFC та CBO без такого перетворення можуть досягати або перевищувати 1000, що виходить за межі діапазонів, для яких була апробована наведена модель. Це обмежує її застосовність для аналізу великих Java-застосунків та ускладнює порівняння результатів між різними проєктами.

Таким чином, аналіз існуючих досліджень свідчить про доцільність реалізації математичної моделі для оцінювання складності об'єктно-орієнтованого проєктування Java-застосунків, яка поєднає переваги використання багатовимірного статистичного аналізу, нормалізуючих перетворень та відносного представлення метрик RFC і CBO. Це дозволить виявити архітектурні аномалії та адаптувати модель до особливостей реальних програмних проєктів.

Постановка завдання. Метою роботи є підвищення достовірності оцінювання складності об'єктно-орієнтованого проєктування Java-застосунків за рахунок реалізації математичної моделі, яка використовує ключові метрики ПЗ та враховує взаємозалежності між ними, що забезпечить більш точне прогнозування архітектурної складності та своєчасне виявлення аномалій проєктування.

Для досягнення поставленої мети необхідно вирішити такі задачі:

– проаналізувати наявні підходи та методи оцінювання складності об'єктно-орієнтованого ПЗ;

– дослідити основні метрики Java-застосунків, що характеризують зв'язність, відповідальність класів та структурну складність ПЗ;

– визначити взаємозалежності між обраними метриками та їх вплив на загальний рівень складності проєктування ПЗ;

– розробити математичну модель для оцінювання складності об'єктно-орієнтованого проєктування на основі багатофакторного аналізу та побудови еліпсів прогнозування;

– реалізувати програму для оцінювання складності Java-застосунків на основі розробленої моделі та перевірити її ефективність на практичних прикладах.

Виклад основного матеріалу. У процесі оцінювання складності об'єктно-орієнтованого проєктування Java-застосунків важливим етапом є перевірка статистичних властивостей вихідних даних. Більшість методів багатовимірного статистичного аналізу, зокрема побудова еліпсів прогнозування та використання відстані Махаланобіса, ґрунтуються на припущенні багатовимірної нормальності розподілу досліджуваних даних. Порушення цього припущення може призводити до викривлення результатів і помилкової класифікації складності проєктування програмних проєктів.

У межах даного дослідження для аналізу складності архітектури використовуються метрики RFC та CBO, які перед побудовою математичної моделі перетворюються у відносний вигляд шляхом ділення на кількість класів NCL. Таке перетворення дозволяє зменшити вплив розміру проєкту та забезпечити можливість порівнювати результати різних проєктів. Отже, подальший аналіз здійснюється для показників RFC/NCL та CBO/NCL. Ці метрики можуть бути визначені вже на ранньому етапі планування проєкту на основі концептуальної моделі застосунку.

У даному дослідженні розглядається набір даних із 302 значень. Набір даних метрик коду для 140 Java-застосунків ігрових рушіїв, розміщених на платформі GitHub [9], було зібрано авторами. Ці дані були отримані за допомогою статичного аналізу вихідного коду з використанням програми SourceMeter for Java [10]. Метрики 88 проєктів було взято з джерела [11] та 74 проєктів – із джерела [12].

Для перевірки відповідності двовимірного набору даних припущенню багатовимірної нормальності застосовано критерій Мардіа, який базується на розрахунку багатовимірної асиметрії та ексцесу [13]. Обчислення показників β_1 та β_2 дозволяє кількісно оцінити ступінь відхилення емпіричного розподілу від гаусівського. Якщо хоча б один із критеріїв перевищує відповідні критичні значення, припущення про багатовимірну нормальність вважається порушеним, що унеможливлює безпосереднє застосування класичних статистичних моделей.

Значення багатовимірної асиметрії β_1 становить 3,58, а відповідна тестова статистика дорівнює 180,61, що значно перевищує критичне значення χ^2-розподілу 14,86 при чотирьох ступенях свободи та рівні значущості 0,005. Це свідчить про наявність суттєвої асиметрії розподілу та не дозволяє вважати його багатовимірно нормальним за критерієм асиметрії.

Значення багатовимірного ексцесу β_2 дорівнює 13,86, а відповідна тестова статистика становить 12,75, що також перевищує критичне значення 2,57. Отриманий результат вказує на надмірну піковість або плоскість розподілу та підтверджує відхилення від багатовимірної нормальності.

У зв'язку з цим для нормалізації даних було застосовано багатовимірне нормалізуюче перетворення Бокса–Кокса, яке вимагає строго додатних значень усіх змінних [14]:

$$Z = \begin{cases} \dfrac{X^{\lambda} - 1}{\lambda}, if\ \lambda \neq 0; \\ \log(X), if\ \lambda = 0, \end{cases} \quad (1)$$

де X приймає значення відповідних стовпців даних RFC/NCL та CBO/NCL.

Оптимальний вектор параметрів $\lambda=[\lambda_1, \lambda_2]$ оцінюється за допомогою методу максимальної правдоподібності (MLE), причому функція правдоподібності включає логарифм визначника коваріаційної матриці нормалізованих даних та перетворення Якобі [15]:

$$L(\lambda) = -\frac{n}{2} log * det\left(Cov(Z)\right) - \sum_{j=1}^{p}\left(\lambda_j - 1\right)\sum_{i=1}^{n} log\left(X_{ij}\right). \quad (2)$$

Отримано вектор параметрів $\lambda = [0{,}276, 0{,}227]$.

У процесі нормалізації здійснювався ітеративний пошук і видалення викидів із використанням квадрату відстані Махаланобіса та порогового значення статистичного тесту, побудованого на основі *F*-розподілу. Розрахунок та видалення викидів було описано у [16, 17]. Загалом було видалено 4 точки даних. Описова статистика очищеного набору даних приведена у таблиці 1.

У межах запропонованої моделі було побудовано два еліпси прогнозування з різними рівнями значущості для нормалізованих та вихідних значень метрик RFC/NCL та CBO/NCL, що дозволяє розділити простір метрик на зони з різним рівнем складності архітектури.

Перший еліпс відповідає рівню значущості α = 0,05 і визначає область типової або низької складності (Low). Значення метрик, що потрапляють усередину цього еліпса, характеризують проєкти з архітектурою, близькою до середньостатистичної, без ознак надмірної складності.

Другий еліпс, побудований для рівня значущості α = 0,005, визначає межу високої складності (High). Простір між двома еліпсами відповідає граничній зоні, в якій архітектура проєкту може вважатися потенційно складною. Дані, що знаходяться поза межами другого еліпса (Out of range), розглядаються як такі, що не можуть бути коректно інтерпретовані в межах даної моделі.

Застосування двох еліпсів дозволяє не лише виявляти викиди, але й більш гнучко класифікувати проєкти за рівнем складності, виділяючи нормальну, граничну та аномальну зони значень метрик.

Рівняння для нормалізованого еліпса прогнозування, побудоване в рамках даної роботи:

$$\frac{(Z_1-1{,}98)^2}{0{,}82}+\frac{(Z_2-2{,}82)^2}{0{,}31}-\frac{0{,}64(Z_1-1{,}98)(Z_2-2{,}82)}{0{,}32}=6{,}54.$$

Рівняння для трансформованого еліпса прогнозування, побудоване в рамках даної роботи:

$$\frac{(\psi_1(X_1)-1{,}98)^2}{0{,}82}+\frac{(\psi_2(X_2)-2{,}82)^2}{0{,}31}-$$
$$-\frac{0{,}64(\psi_1(X_1)-1{,}98)(\psi_2(X_2)-2{,}82)}{0{,}32}=6{,}54.$$

Для перевірки працездатності побудованої моделі було проаналізовано 26 Java-проєктів, які використовувалися як тестова вибірка. Дані проєкти були взяті з відкритого репозитарію GitHub [9]. Для кожного проєкту виконувалася нормалізація значень метрик RFC/NCL та CBO/NCL і визначався квадрат відстані Махаланобіса від центра еліпсів прогнозування. Результати перевірки моделі та визначення складності проєктування наведено в таблиці 2. Побудовані еліпси прогнозування для нормалізованих даних та трансформовані еліпси прогнозування для вихідних даних наведено на рисунках 1-2.

Модель може бути застосована для проєктів, значення метрик яких потрапляють у такі межі: NCL від 4 до 7874, RFC від 3 до 37,278 та CBO від 0,696 до 22,417.

Аналіз отриманих результатів для тестової вибірки показав, що більшість проєктів мають низький рівень складності об'єктно-орієнтованого проєктування. Проєкти, віднесені до категорій High, мають значення окремих метрик, які суттєво відрізняються від середніх показників, що може вказувати на надмірну складність або наявність нетипових структур. Проєкти, віднесені до категорій Out of range, не можуть бути адекватно описані побудованою моделлю та потребують побудови окремої моделі для їх аналізу.

Таким чином, побудована математична модель на основі двох еліпсів прогнозування дозволяє ефективно оцінювати складність об'єктно-орієнтованого проєктування Java-застосунків, розрізняти їх за типами та виявляти аномальні проєкти. Для автоматизації розрахунків практична реалізація побудованої моделі була виконана у відповідній програмі та представлена у [18].

Висновки. У роботі побудовано математичну модель для оцінювання складності об'єктно-орієнтованого проєктування Java-застосунків на основі метрик RFC та CBO. Запропонований підхід враховує взаємозалежність метрик ПЗ та спрямований на підвищення достовірності виявлення архітектурних аномалій.

У ході дослідження було розглянуто сучасні методи оцінювання архітектури ПЗ та визначено,

Таблиця 1

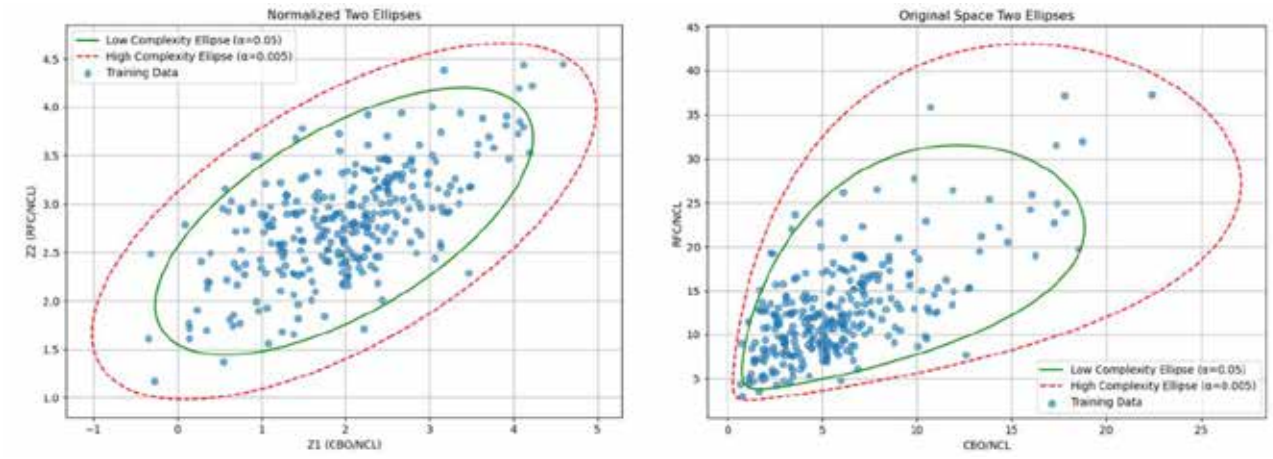
Описова статистика очищеного набору даних (*N*=298)

Назва метрики	Мінімальне	Максимальне	Середнє	Середньоквадратичне відхилення
NCL	4	7874	547,161	1025,952
CBO/NCL	0,696	22,417	5,865	3,678
RFC/NCL	3,000	37,278	12,636	5,609

Таблиця 2

Тестові емпіричні дані із оцінюванням складності проєктування

Назва	CBO	RFC	NCL	CBO/NCL	RFC/NCL	Conclusion
bookstore	92	210	26	3,53	8,07	Low
calculator-with-android-studio-Java	1	9	4	0,25	2,25	Out of range
cel-java	4947	12514	774	6,39	16,16	Low
Chat-App	84	331	52	1,61	6,36	Low
chess-ai	168	412	27	6,22	15,25	Low
glide-transformations	25	175	25	1	7	Low
glygen.cfde.generator	209	918	80	2,61	11,47	Low
java-tor	75	250	31	2,41	8,06	Low
JavaAsciiGenerator	1	15	2	0,5	7,5	High
JavaSudokuSolver	2	32	3	0,66	10,66	High
kroxylicious	3262	9936	1006	3,24	9,87	Low
logger	48	149	14	3,42	10,64	Low
logstash	2894	7391	737	3,92	10,02	Low
lombok	8438	15491	3604	2,34	4,29	Low
mcav	1017	2385	305	3,33	7,81	Low
misle-java	611	1622	94	6,5	17,25	Low
mockito	5988	12122	1825	3,28	6,64	Low
plantuml	30971	56076	3475	8,91	16,13	Low
presto	5182	16228	1186	4,36	13,68	Low
r2cloud	2319	7452	453	5,11	16,45	Low
recipe-app-java	30	51	14	2,14	3,64	High
Synapse	23	75	16	1,43	4,68	Low
TaskTracker	39	118	16	2,43	7,37	Low
TextAnalyzr	1	19	2	0,5	9,5	Out of range
webmagic	1130	2614	309	3,65	8,45	Low
yp-online-store-showcase	270	359	79	3,41	4,54	Low



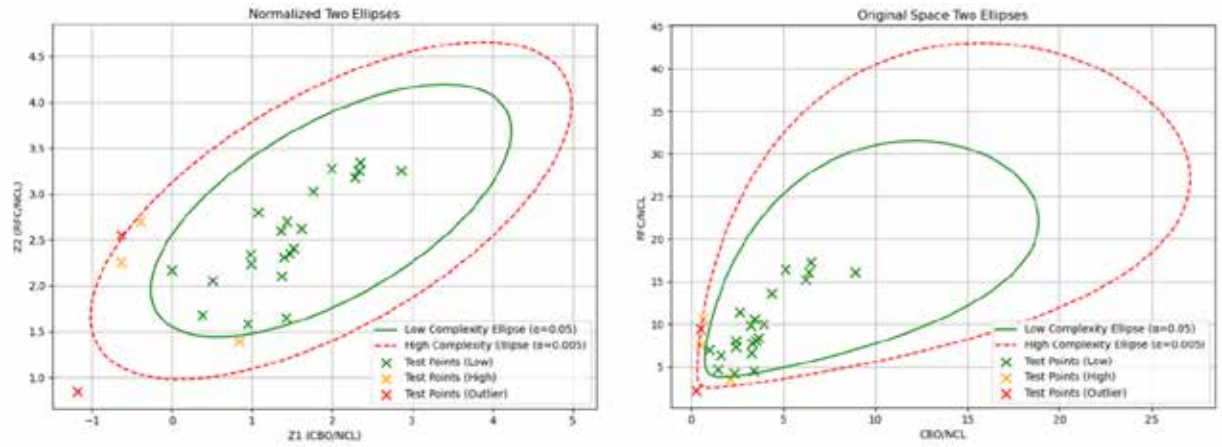
а) б)

Рис. 1. Еліпси прогнозування для тренувальних даних (*N*=298)

а) еліпси прогнозування для нормалізованих даних; б) трансформовані еліпси прогнозування для вихідних даних

що метрики RFC і CBO найкраще відображають складність структури й взаємодію класів у Java-застосунках. Було зібрано та проаналізовано дані з відкритих джерел, на основі яких сформовано вибірки для побудови та перевірки моделі.

Побудовану модель подано у вигляді двох еліпсів прогнозування. Для підвищення достовірності оцінювання використано нормалізацію за багато-



а) б)

Рис. 2. Еліпси прогнозування для тестових даних (*N*=26)

а) еліпси прогнозування для нормалізованих даних; б) трансформовані еліпси прогнозування для вихідних даних

вимірним перетворенням Бокса–Кокса, викиди видалялися за допомогою квадрата відстані Махаланобіса.

Загалом, побудована модель дозволяє адекватно оцінювати складність архітектурних рішень і може бути використана для аналізу складності об'єктно-орієнтованого проєктування Java-застосунків. Результати тестування підтвердили її адекватність і здатність коректно виявляти аномальні проєкти. У межах роботи також реалізовано програму для автоматизації оцінювання складності об'єктно-орієнтованого проєктування Java-застосунків, практичну придатність якої підтверджено експериментально.

Список літератури:

1. Chidamber S.R., Kemerer C.F. Towards a metrics suite for object oriented design. *ACM SIGPLAN Notices.* 1991. Vol. 26, no. 11. P. 197–211. DOI: https://doi.org/10.1145/118014.117970

2. Suresh Y., Pati J., Rath S. Ku. Effectiveness of software metrics for object-oriented system. *Procedia Technology*. 2012. Vol. 6. P. 420–427. DOI: https://doi.org/10.1016/j.protcy.2012.10.050

3. Software Metrics. URL: https://www.imagix.com/links/software-metrics.html (дата звернення: 12.01.26).

4. Khan S.A., Khan R.A. Object oriented design complexity quantification model. *Procedia Technology*. 2012. Vol. 4. P. 548–554. DOI: https://doi.org/10.1016/j.protcy.2012.05.087

5. Prykhodko S., Makarova L., Prykhodko K., Pukhalevych A. Application of Transformed Prediction Ellipsoids for Outlier Detection in Multivariate Non-Gaussian Data. *Proceedings of the IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*. 2020. P. 359–362. DOI: https://doi.org/10.1109/TCSET49122.2020.235454

6. Johnson R.A., Wichern D.W. Applied multivariate statistical analysis. Pearson Prentice Hall, 2007. 800 p.

7. Komolov S., Dlamini G., Megha S., Mazzara M. Towards Predicting Architectural Design Patterns: A Machine Learning Approach. *Computers*. 2022. Vol. 11, no. 10, art. 151. DOI: https://doi.org/10.3390/computers11100151

8. Prykhodko S., Prykhodko N., Smykodub T. A Joint Statistical Estimation of the RFC and CBO Metrics for Open-Source Applications Developed in Java. *Proceedings of the IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT).* 2022. P. 160–166. DOI: https://doi.org/10.1109/CSIT56902.2022.10000457

9. GitHub. URL: https://github.com (дата звернення: 15.01.2026).

10. SourceMeter for Java. URL: https://sourcemeter.com (дата звернення: 15.01.2026).

11. Prykhodko S., Makarova L., Pukhalevych A. Statistical Analysis of the Three-Dimensional Data of Software Metrics RFC, CBO, and WMC that are not Normally Distributed. *Proceedings of International Conference on Applied Innovation in IT.* 2025. Vol. 13, issue 1. P. 127–132. DOI: https://doi.org/10.25673/119254

12. Oriekhov O., Farionova T., Chernova L.S., Chernova L., Vorona M. Nonlinear regression models for software size estimation of Data Science and Machine Learning Java-applications. *Proceedings of the 5th International Workshop IT Project Management (ITPM).* 2024. CEUR Workshop Proceedings. 2024. Vol. 3709. P. 54–66. DOI: https://doi.org/10.23939/IW_itpm2024.054

13. Mardia K.V. Measures of multivariate skewness and kurtosis with applications. *Biometrika*. 1970. Vol. 57. P. 519–530. DOI: https://doi.org/10.1093/biomet/57.3.519

14. Using Box-Cox to Achieve Multivariate Normality. URL: https://richard-mei97.medium.com/using-box-cox-to-achieve-multivariate-normality-3957f0164e8f (дата звернення: 15.01.2026).

15. Maximum Likelihood Method. URL: https://randomservices.org/random/prob/index.html (дата звернення: 15.01.2026).

16. Prykhodko S., Makarova L., Latanska L., Bryzghalov M. Mathematical Model for Detecting Outliers in the Two-Dimensional Data of Software Metrics RFC and CBO from Applications in Java. *Proceedings of the 13th International Conference on Information Control Systems & Technologies (ICST).* 2025. CEUR. Vol. 4048. P. 509–519. URL: https://ceur-ws.org/Vol-4048/paper40.pdf (дата звернення: 15.01.2026).

17. Приходько С., Макарова Л., Латанська Л., Бризгалов М. Виявлення викидів у двовимірних даних програмних метрик RFC та CBO для Java застосунків з використанням трансформованого еліпсу прогнозування. *Інформаційні управляючі системи і технології (ІУСТ-ОДЕСА-2025): матеріали XIII Міжнародної науково-практичної конференціі (24–26 вересень 2025 р., Одеса)* / вип. ред. В.В. Вичужанін, 2025. С. 244–247.

18. Бризгалов М.В., Макарова Л.М. Розробка програми для оцінювання складності об'єктно-орієнтованого проєктування Java застосунків з використанням метрик RFC та CBO. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2025): Матеріали VI Міжнародної науково-практичної інтернет конференції (16-17 листопада 2025 р.)*. Миколаїв: НУК імені адмірала Макарова, 2025. С. 270–272.

Makarova L.M., Pukhalevych A.V., Bryzgalov M.V., Kaminsky S.S. MATHEMATICAL MODELING FOR ASSESSING THE COMPLEXITY OF OBJECT-ORIENTED DESIGN OF JAVA APPLICATIONS USING RFC AND CBO METRICS

This paper addresses the problem of assessing the complexity of object-oriented design of Java applications based on the software metrics RFC (Response for a Class) and CBO (Coupling Between Objects). The relevance of this problem is driven by the need for timely detection of architectural anomalies and excessive software complexity at early stages of development, since such deviations significantly affect the quality, maintainability, and cost of software products. Existing approaches to design complexity analysis often fail to provide sufficient reliability of results and do not take into account the statistical characteristics of metric value distributions, which can lead to incorrect conclusions about the state of an application's architecture.

To address this issue, the paper proposes a mathematical model for assessing the complexity of object-oriented design of Java applications based on the construction of a prediction ellipse in the space of RFC and CBO metrics. To improve the reliability of the assessment, a multivariate Box–Cox transformation is applied to normalize the statistical data, along with multivariate statistical analysis methods, in particular the use of the Mahalanobis distance for outlier detection. The proposed approach enables more accurate identification of applications with anomalous architectural complexity and reduces the impact of noise in the input data on the analysis results.

Based on the developed mathematical model, software for automated assessment of the complexity of Java applications has been created. Testing of the software confirmed the effectiveness of the proposed approach and its advantages over traditional metric analysis methods, particularly in terms of outlier detection and improving the reliability of object-oriented design complexity assessment.

The obtained results can be used by developers and software project managers to control the architectural complexity of Java applications, optimize design and maintenance processes, and reduce risks associated with increasing complexity and declining software quality.

Keywords: *design complexity, Java application, RFC and CBO metrics, prediction ellipse, Box–Cox transformation, Mahalanobis distance, outlier detection.*

Дата першого надходження статті до видання: 27.01.2026
Дата прийняття статті до друку після рецензування: 24.02.2026
Дата публікації (оприлюднення) статті: 08.04.2026