



О. Б. Керницький, Ю. В. Опотяк, В. М. Теслюк

Національний Університет "Львівська політехніка", м. Львів, Україна

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ СИНТЕЗУ ТА ВАЛІДАЦІЇ СПЕЦИФІКАЦІЙ ВИМОГ ІТ-ПРОЄКТІВ У ЗАДАЧАХ РЕІНЖІНІРИНГУ LEGACY-СИСТЕМ

Перебудова застарілих інформаційних систем потребує не тільки технічного оновлення програмного забезпечення, а й відновлення, уточнення та перевірки специфікацій вимог до ІТ-проектів, які мають відображати реальну поведінку системи. Можливим вирішенням цієї проблеми є розроблення інформаційної технології синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем. Актуальність дослідження зумовлена необхідністю формування специфікацій вимог до ІТ-проектів, що коректно відображають бізнес-логіку системи, її функціональні залежності та поведінку в умовах модернізації. Особливістю розробленої інформаційної технології є поєднання декомпозиції та аналізу legacy-систем, синтезу вимог на основі реверс-інжинірингу, формалізації бізнес-процесів на основі мереж Петрі та статистичної валідації специфікацій ІТ-проектів за меж єдиного ітеративного циклу. У роботі розроблено методологію синтезу та валідації специфікацій ІТ-проектів, визначено її методологічні принципи, сформовано структуру інформаційної технології практичної реалізації цієї методики, розроблено сукупність взаємопов'язаних методів і алгоритмів, а також виконано програмну реалізацію у вигляді експериментального кроссплатформного програмного комплексу. Для окремих складових технології наведено математичні моделі у вигляді кортежів, які описують декомпозицію та аналіз legacy-систем, синтез вимог, формалізацію бізнес-процесів і статистичну валідацію специфікацій ІТ-проектів. Практичну придатність розробленої інформаційної технології показано на прикладі специфікації продукту Consumer Credit Report (SRS v1.1.37), зокрема – для валідації вхідних запитів, бізнес-логіки оброблення та формування вихідного звіту. Вимоги та правила валідації сформовано на основі реверс-інжинірингу, валідовано шляхом паралельного тестування, формалізовано у вигляді мережі Петрі та перевірено засобами стохастичної симуляції поведінки системи. Експериментальні результати підтвердили працездатність розробленої інформаційної технології та показали, що застосування статистичної валідації дає змогу зменшити обчислювальні витрати приблизно в 1,5–3,0 рази порівняно з повною формальною перевіркою зі збереженням потрібного рівня достовірності результатів. Отримані результати підтвердили доцільність використання розробленої інформаційної технології для підтримки процесів реінжинірингу legacy-систем, уточнення специфікацій ІТ-проектів і контролю їх узгодженості з реальною поведінкою системи. Результати дослідження можуть бути використані під час модернізації та реінжинірингу складних legacy-систем.

Ключові слова: реверс-інжиніринг; синтез вимог; аналіз специфікацій ІТ-проектів; паралельне тестування; модель чорного ящика; матриця відстеження вимог; аналіз функціональних розбіжностей; мережі Петрі; статистична валідація специфікацій вимог до ІТ-проектів.

Вступ / Introduction

Реінжиніринг legacy-систем є одним із найпоширеніших напрямів модернізації сучасних ІТ-рішень, оскільки значна кількість прикладних систем продовжує функціонувати за умов застарілої архітектури, неповної документації та високої залежності від історично сформованої бізнес-логіки. За таких умов модернізація системи потребує не тільки технічного оновлення програмного забезпечення, а й відновлення, уточнення та перевірки специфікацій, які мають відображати її реальну поведінку. Саме якість специфікацій вимог до ІТ-проектів на ранніх кроках реінжинірингу legacy-систем значною мірою визначає коректність подальшого проектування, зменшує ризик функціональних відхи-

лень і дає змогу забезпечити узгодженість нової системи з бізнес-потребами та поведінкою наявного рішення.

Особливої складності ця задача набуває у випадках, коли замовник вимагає, щоб нова система зберігала функціональну відповідність наявній системі, але, при цьому, була реалізована на сучасній технологічній основі. За таких умов виникає потреба у відокремленні вимог, зумовлених обмеженнями архаїчної системи, від вимог, які дійсно відображають бізнес-потреби, що потребує поєднання поведінкового аналізу, формалізації бізнес-процесів і масштабованої перевірки коректності специфікацій вимог до ІТ-проектів.

У роботі розроблено інформаційну технологію, яка поєднує декомпозицію та аналіз legacy-систем, синтез вимог на основі реверс-інжинірингу, формалізацію біз-

© Copyright 2026 by the author(s)

Керницький Олег Богданович, аспірант, кафедра автоматизованих систем управління.

Email: oleh.b.kernytskyi@lpnu.ua; <https://orcid.org/0009-0007-5318-6506>

Опотяк Юрій Володимирович, канд. техн. наук, ст. науковий співробітник, кафедра автоматизованих систем управління.

Email: yurii.v.opotyak@lpnu.ua; <https://orcid.org/0000-0001-9889-4177>

Теслюк Василь Миколайович, д-р техн. наук, професор, завідувач кафедри автоматизованих систем управління.

Email: vasyl.m.teslyuk@lpnu.ua; <https://orcid.org/0000-0002-5974-9310>

Цитування за ДСТУ: Керницький О. Б., Опотяк Ю. В., Теслюк В. М. Інформаційна технологія синтезу та валідації специфікацій вимог ІТ-проектів у задачах реінжинірингу legacy-систем. Науковий вісник НЛТУ України. 2026, т. 36, № 2. С. 168–180.

Citation APA: Kernytskyi, O. B., Opotyak, Yu. V., & Teslyuk, V. M. (2026). Information technology of specifications synthesis and validation in the process of it project reengineering. *Scientific Bulletin of UNFU*, 36(2), 168–180. <https://doi.org/10.36930/40360219>

нес-процесів на основі мереж Петрі та статистичну валідацію специфікацій ІТ-проектів. У подальших розділах статті послідовно розкрито постановку проблеми й обмеження наявних підходів, методологічні принципи та місце інформаційної технології, її методи, алгоритми реалізації, програмну реалізацію та результати експериментального дослідження, яке, зокрема, підтверджує зменшення обчислювальних витрат в 1,5-3,0 рази зі збереженням потрібного рівня достовірності результатів.

Об'єкт дослідження – синтез та валідація специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем.

Предмет дослідження – методи і засоби синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем, що дасть змогу визначити їх місце у задачах реверс-інжинірингу, функціональне призначення та можливості інтеграції в межах єдиної інформаційної технології.

Мета роботи – розробити методику синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем та інформаційну технологію її практичної реалізації, що дасть змогу поєднати поведінковий аналіз, формальне моделювання та масштабовану статистичну перевірку синтезованих специфікацій ІТ-проектів.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- проаналізувати обмеження традиційних підходів до синтезу специфікацій, пов'язаних з ІТ-проектами перебудови legacy-систем, що дасть змогу виявити їх недостатню придатність для відновлення, уточнення та перевірки специфікацій вимог до ІТ-проектів у складних умовах реінжинірингу legacy-систем;
- проаналізувати наявні методи та алгоритми, які забезпечують синтез та валідацію специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем, що дасть змогу визначити їх місце у задачах реверс-інжинірингу, функціональне призначення та можливості інтеграції в межах єдиної інформаційної технології;
- удосконалити методи декомпозиції та аналізу legacy-систем задля застосування їх у розробленні інформаційної технології синтезу та валідації специфікацій вимог до ІТ-проектів, що допоможе підвищити повноту, узгодженість і формалізованість синтезованих специфікацій вимог до ІТ-проектів;
- розробити методику синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем, яка об'єднає методи декомпозиції та аналізу legacy-систем, синтезу вимог, формалізації бізнес-процесів і статистичної валідації специфікацій, що забезпечить цілісний ітеративний процес аналізу, формалізації, перевірки та уточнення специфікацій;
- розробити інформаційну технологію практичної реалізації запропонованої методики, яка забезпечить алгоритмічну, програмну та інструментальну підтримку синтезу й валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем.

Аналіз останніх досліджень та публікацій. У низці сучасних наукових праць дослідники розглянули різні аспекти модернізації legacy-систем, реверс-інжинірингу, формалізації бізнес-процесів та статистичної перевірки моделей. Зокрема, у роботі [7] висвітлено проблематику міграції legacy-систем до хмарного середовища з архітектурної точки зору. Автори проаналізували ключові архітектурні труднощі такої міграції, підкреслено значення попереднього дослідження наявної сис-

теми та наголошено на потребі врахування залежностей між компонентами, що безпосередньо пов'язано із відновленням та уточненням специфікацій вимог до ІТ-проектів у процесі реінжинірингу legacy-систем.

У праці [1] запропоновано таксономію підходів до ідентифікації сервісів у процесі модернізації legacy-систем. Автори систематизують підходи до виділення функціональних сервісів та описують їх роль у перетворенні застарілих систем у більш гнучкі архітектури. Це свідчить про важливість попередньої декомпозиції системи та структуризації її функціональних складових як окремого кроку модернізації.

У роботі [6] узагальнено сучасні стратегії модернізації програмного забезпечення, чинники, що визначають доцільність модернізації, а також перспективні напрями подальших досліджень у цій сфері. Автори наголошують, що модернізація систем не може зводитися тільки до технічного перенесення функціональності, а потребує чіткого розуміння логіки роботи наявної системи, її архітектурних обмежень та способів формалізованого відтворення її поведінки. Подібна думка простежується і в роботі [24], де розглянуто системно-інженерні бар'єри еволюції legacy-систем та наголошено на складності оцінювання їх поточного стану, залежностей і ризиків під час модернізації.

Окремий напрям становлять роботи з реверс-інжинірингу як засобу відновлення знань про наявну систему. У роботі [23] подано систематичний огляд model-driven підходів до реверс-інжинірингу. Автори аналізують моделі, модельні трансформації та програмні засоби, які застосовують для відновлення архітектури і поведінки систем, та показують, що модельно-орієнтовані підходи є перспективними для подальшої формалізації процесів. Своєю чергою, у роботі [14] запропоновано підхід Retriever до реверс-інжинірингу архітектурних моделей програмного забезпечення. Автори зосереджують увагу на поданні системи через множини взаємопов'язаних представлень, що є важливим з точки зору відновлення структурних і поведінкових аспектів legacy-систем. У роботі [18] досліджено reverse-MBSE підхід до legacy-систем на прикладі конкретного кейсу. Робота підтверджує доцільність застосування реверс-інжинірингу не тільки для відновлення артефактів системи, а й для подальшого переходу до модельно-орієнтованого опису.

Істотне значення для завдань синтезу специфікацій вимог до ІТ-проектів мають праці з формалізації бізнес-процесів. У роботі [17] розглянуто підхід до моделювання і симуляції бізнес-процесів на основі мереж Петрі. Автори показали, що застосування мереж Петрі дає змогу подати процеси у строгому формальному вигляді, описати послідовні, паралельні та альтернативні гілки виконання і створити основу для аналізу поведінки системи. Такий підхід є важливим для усунення неоднозначностей текстових специфікацій та переходу до їх перевірки на формальному рівні.

Оскільки повна формальна перевірка складних моделей супроводжується значними обчислювальними витратами, окрему групу становлять роботи зі статистичної перевірки моделей. У праці [4] досліджено statistical model checking для стохастичних timed-arc мереж Петрі. Автори розглядають застосування статистичних процедур до формальних моделей і показують можливість оцінювання їх властивостей без повного перебору всіх станів. Близький за змістом підхід подано в роботі

[2], де досліджено sound statistical model checking для ймовірностей та очікуваних винагород. Ці результати підтверджують перспективність використання статистичних підходів для перевірки складних моделей у випадках, коли традиційна формальна верифікація є надто ресурсомісткою.

Отже, проведений аналіз сучасних літературних джерел показав, що в наукових працях достатньо широко висвітлено окремі аспекти модернізації legacy-систем, реверс-інжинірингу, формалізації процесів та статистичної перевірки моделей. Однак ці підходи здебільшого розглядають ізольовано: одні праці зосереджені на архітектурній модернізації та декомпозиції систем, інші – на модельно-орієнтованому реверс-інжинірингу, формальному поданні бізнес-процесів або статистичній перевірці моделей. З аналізу публікацій впливає актуальність завдання підвищення ефективності синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем, що потребує розроблення цілісної інформаційної технології, яка б поєднувала декомпозицію системи, синтез вимог, формалізацію бізнес-процесів і масштабовану статистичну валідацію специфікацій вимог до ІТ-проектів.

Матеріали та методи дослідження. У дослідженні використано матеріали, що характеризують функціонування legacy-систем, сценарії їх роботи, наявні специфікації та результати експериментального застосування розробленої інформаційної технології. У роботі застосовано методи декомпозиції та аналізу legacy-систем, метод синтезу вимог на основі реверс-інжинірингу, метод формалізації бізнес-процесів на основі мереж Петрі та метод статистичної валідації специфікацій вимог до ІТ-проектів, які формують методологічну основу інформаційної технології синтезу та валідації специфікацій вимог до ІТ-проектів, орієнтованої на задачі реінжинірингу legacy-систем. Розроблена методологія синтезу та валідації специфікацій вимог до ІТ-проектів ґрунтується на сукупності принципів: поведінкової відповідності, інтеграції реверс-інжинірингу та формального моделювання, формалізації процесів, масштабованої валідації, ітеративного уточнення специфікацій вимог до ІТ-проектів та інтегрованості підсистем. У процесі застосування інформаційної технології реалізується замкнений цикл, у якому результати перевірки використовують для уточнення вимог і повторної формалізації процесів, що забезпечує поступове підвищення повноти специфікацій вимог до ІТ-проектів для узгодженості з поведінкою legacy-системи. Для аналізу поведінки систем використано black-box підхід (модель чорного ящика) і паралельне тестування, а для перевірки специфікацій вимог до ІТ-проектів – формальне моделювання та стохастичний аналіз трас виконання. Практичну апробацію виконано з використанням експериментального програмного комплексу.

Результати дослідження та їх обговорення / Research results and their discussion

Аналіз обмежень традиційних підходів до синтезу специфікацій вимог до ІТ-проектів. Реінжиніринг legacy-систем супроводжується необхідністю відновлення, уточнення та перевірки специфікацій вимог до ІТ-проектів, які здебільшого або відсутні, або не відображають фактичної логіки функціонування системи [6, 8, 24]. У процесі тривалої експлуатації такі системи заз-

нають численних модифікацій, що призводить до накопичення локальних змін, втрати цілісності документації та розходження між задекларованими вимогами і реальною поведінкою програмного забезпечення. Унаслідок цього модернізація системи без достовірно відновленої специфікації супроводжується високим ризиком функціональних помилок, порушення бізнес-правил і повторного виконання частини робіт.

Проблема ускладнюється тим, що відновлення специфікацій вимог до ІТ-проектів не зводиться тільки до аналізу програмного коду або наявної документації. У багатьох випадках вирішальне значення має саме поведінковий аналіз системи, оскільки значна частина фактичної логіки реалізована через історично сформовані правила оброблення даних, неформалізовані сценарії виконання та приховані залежності між підсистемами. Це означає, що процес реінжинірингу legacy-систем потребує не тільки виявлення функціональних компонентів системи, а й відновлення причинно-наслідкових зв'язків між вхідними даними, станами процесів і результатами виконання.

Наявні підходи до вирішення цієї проблеми мають низку істотних обмежень. Традиційний реверс-інжиніринг переважно зосереджений на відновленні структури системи, аналізі вихідного коду або залежностей між модулями, однак не завжди дає змогу відтворити повну поведінкову модель системи на рівні бізнес-процесів [14, 18, 23]. Підходи, що ґрунтуються тільки на аналізі документації, не забезпечують достатньої достовірності, оскільки документація legacy-систем часто є неповною, застарілою або суперечливою. Black-box аналіз, навпаки, дає змогу досліджувати фактичну поведінку системи, однак без формалізації результатів такого аналізу не забезпечує строгого подання специфікацій вимог до ІТ-проектів та їх подальшої перевірки.

Формальні методи моделювання, зокрема – підходи на основі мереж Петрі, дають змогу усунути неоднозначність текстових описів та подати бізнес-процеси у строгому вигляді [9, 11, 17]. Проте їх самостійне застосування також має обмеження. Побудова формальної моделі без попереднього відновлення та уточнення вимог не гарантує відповідності моделі реальній поведінці системи. Окрім цього, повна формальна перевірка складних моделей супроводжується швидким зростанням кількості станів, що істотно ускладнює застосування таких підходів до великих і багатокomпонентних legacy-систем.

Статистичні методи аналізу моделей частково усувають проблему обчислювальної складності, оскільки дають змогу перейти від повного перебору станів до вибіркового аналізу поведінки [2, 4, 12]. Однак їх використання поза контекстом цілісної технології також є недостатнім, оскільки без попередніх кроків декомпозиції системи, синтезу вимог і формалізації процесів статистична перевірка не може забезпечити належної якості специфікацій вимог до ІТ-проектів.

Отже, науково-практична проблема полягає у відсутності цілісного підходу, який поєднував би декомпозицію legacy-систем, поведінковий аналіз, синтез вимог, формалізацію бізнес-процесів і масштабовану перевірку специфікацій вимог до ІТ-проектів у межах єдиної інформаційної технології. Саме ця обставина зумовила потребу розроблення інформаційної технології, що дає змогу послідовно відновлювати, формалізувати

та валідовувати специфікації ІТ-проектів у задачах реінжинірингу legacy-систем.

Структура інформаційної технології синтезу та валідації специфікацій вимог до ІТ-проектів. Розроблена інформаційна технологія синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем є інтегрованою сукупністю взаємопов'язаних процесів, методів і засобів, що забезпечують відновлення, формалізацію та перевірку специфікацій вимог до ІТ-проектів на основі аналізу реальної поведінки системи [10, 11, 12].

Методологічну основу інформаційної технології можна подати у вигляді кортежу:

$$IT = \langle M_{dec}, M_{syn}, M_{form}, M_{val}, \Psi \rangle, \quad (1)$$

де: M_{dec} – метод декомпозиції та аналізу legacy-систем; M_{syn} – метод синтезу вимог на основі реверс-інжинірингу; M_{form} – метод формалізації бізнес-процесів на основі мереж Петрі; M_{val} – метод статистичної валідації специфікацій вимог до ІТ-проектів; Ψ – механізм ітеративного уточнення специфікацій вимог до ІТ-проектів за результатами перевірки.

Інформаційна технологія (рис. 1) ґрунтується на принципі покрокової трансформації знань про систему, від неформалізованих спостережень до формальних моделей та їх статистичної валідації і реалізує ітеративний цикл уточнення специфікацій вимог до ІТ-проектів.



Рис. 1. Підсистеми інформаційної технології синтезу та валідації специфікацій вимог до ІТ-проектів під час реінжинірингу legacy-систем / Subsystems of the information technology for developing and validating specifications in IT project re-engineering of legacy systems

Загальна структура інформаційної технології містить три основні функціональні підсистеми, що узгоджується із сучасними підходами до поєднання реверс-інжинірингу, формального моделювання та масштабованої перевірки моделей [4, 6, 23]:

- підсистему реверс-інжинірингу та поведінкового аналізу;
- підсистему формалізації бізнес-процесів;
- підсистему статистичної валідації специфікацій вимог до ІТ-проектів.

Підсистема реверс-інжинірингу та поведінкового аналізу. На початковому кроці у підсистемі реверс-інжинірингу та поведінкового аналізу здійснюється відновлення поведінки legacy-системи на основі аналізу її зовнішніх реакцій. Для цього застосовано black-box підхід, що дає змогу досліджувати систему без доступу до її внутрішньої реалізації [10, 13].

Ключові компоненти підсистеми вирішують такі поставлені завдання:

- формування тестових сценаріїв;
- виконання серій експериментів над системою;
- фіксація вхідних даних і відповідних вихідних реакцій;
- порівняння поведінки наявної та модернізованої систем у режимі паралельного тестування.

Результатом функціонування підсистеми є множина узгоджених сценаріїв виконання та відновлених функціональних залежностей, що відображають реальну логіку роботи системи і слугують основою для формування специфікацій вимог до ІТ-проектів.

Підсистема формалізації бізнес-процесів. У цій підсистемі синтезовані сценарії виконання трансформуються у формальні моделі бізнес-процесів. Для цього використано апарат мереж Петрі, що забезпечує стро-

гий математичний опис динамічної поведінки системи. Формалізація містить:

- відображення станів процесів у вигляді позицій;
- подання подій та операцій у вигляді переходів;
- моделювання потоків керування через систему дуг;
- опис паралельних, альтернативних та синхронізованих процесів.

Формальні моделі дають змогу усунути неоднозначність текстових специфікацій вимог до ІТ-проектів і створюють основу для подальшого аналізу та перевірки властивостей системи.

Підсистема статистичної валідації специфікацій вимог до ІТ-проектів. Наступна підсистема інформаційної технології виконує перевірку коректності сформованих специфікацій вимог до ІТ-проектів із використанням статистичного аналізу моделей. На відміну від повної формальної верифікації, що потребує перебору всіх можливих станів, застосовується підхід вибіркового генерування трас виконання.

Ключові компоненти підсистеми вирішують такі поставлені завдання:

- генерування множини трас виконання на основі стохастичних процедур;
- оцінювання властивостей моделей за вибіркою сценаріїв;
- розрахунок метрик покриття, достовірності та похибки;
- прийняття рішення щодо коректності специфікацій вимог до ІТ-проектів.

Такий підхід дає змогу виконувати перевірку великих моделей за значно менших обчислювальних витрат порівняно з повною перевіркою.

Ітеративний цикл уточнення специфікацій вимог до ІТ-проектів. Принциповою особливістю інформа-

ційної технології є наявність зворотного зв'язку між підсистемами. Результати статистичної валідації використовують для виявлення невідповідностей або неповноти специфікацій вимог до ІТ-проектів після чого виконується повторний цикл:

- уточнення сценаріїв на етапі реверс-інжинірингу;
- оновлення формальних моделей;
- повторну перевірку специфікацій вимог до ІТ-проектів.

Отже, інформаційна технологія реалізує ітеративний процес, у межах якого відбувається поступове підвищення повноти та узгодженості специфікацій вимог до ІТ-проектів із реальною поведінкою системи. Розроблена структура інформаційної технології забезпечує інтеграцію емпіричного аналізу та формальних методів, поєднання детермінованих і стохастичних підходів до перевірки, масштабованість застосування до великих legacy-систем і системну підтримку процесу реінжинірингу ІТ-проектів.

Методологічні принципи формування інформаційної технології та її місце у задачах реверс-інжинірингу. Розроблена методологія синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем ґрунтується на сукупності принципів, що визначають методологічну основу та забезпечують узгоджене поєднання різнорідних підходів до аналізу, формалізації та перевірки специфікацій вимог до ІТ-проектів. Інформаційну технологію розроблено як практичну реалізацію цієї методики.

Принцип поведінкової відповідності. Інформаційна технологія ґрунтується на припущенні, що специфікації ІТ-системи мають відображати її фактичну поведінку. Відновлення вимог здійснюється на основі аналізу зовнішніх реакцій системи, що дає змогу формувати специфікації, узгоджені з реальною логікою функціонування.

Принцип інтеграції реверс-інжинірингу та формального моделювання. Технологія передбачає поєднання емпіричного аналізу поведінки системи та її формального опису. Результати реверс-інжинірингу трансформуються у формальні моделі бізнес-процесів, що забезпечує однозначність специфікацій вимог до ІТ-проектів та можливість їх подальшої перевірки.

Принцип формалізації процесів. Специфікації подають у вигляді формальних моделей, що дає змогу усунути неоднозначність текстових описів і забезпечити строгий опис поведінки системи. Формалізація дає можливість враховувати паралельність виконання процесів, альтернативні сценарії та механізми синхронізації.

Принцип масштабованої валідації. Перевірка специфікацій вимог до ІТ-проектів виконується з використанням статистичних методів, що дає змогу обмежити обчислювальні витрати зі збереженням потрібного рівня достовірності результатів. Застосування вибіркового генерування трас виконання дає можливість здійснювати перевірку моделей, для яких повний перебір станів є практично неможливим.

Принцип ітеративного уточнення специфікацій вимог до ІТ-проектів. Інформаційна технологія реалізує замкнений цикл, у межах якого результати перевірки використовуються для уточнення вимог і повторної формалізації процесів. Такий підхід забезпечує поступове підвищення повноти специфікацій вимог до ІТ-проектів та їх узгодженості з поведінкою системи.

Принцип інтегрованості підсистем. Усі компоненти інформаційної технології функціонують як єдина

система, у межах якої результати кожного етапу використовуються на наступних етапах, що забезпечує цілісність процесу синтезу та валідації специфікацій вимог до ІТ-проектів й унеможливає розрив між аналізом, моделюванням та перевіркою.

Розроблена інформаційна технологія займає проміжне положення між традиційними методами реверс-інжинірингу, що орієнтовані на відновлення структури системи, та формальними методами аналізу, що забезпечують строгий опис процесів, але мають обмеження щодо масштабованості [2, 4, 14, 23]. На відміну від наявних підходів, вказана інформаційна технологія забезпечує поєднання поведінкового аналізу та формального моделювання, інтеграцію детермінованих і стохастичних методів перевірки, можливість застосування формальної валідації до великих legacy-систем і системну підтримку процесу реінжинірингу ІТ-проектів (рис. 2) [3, 6, 5, 13].



Рис. 2. Схематичне зображення системних переваг забезпечення інформаційною технологією синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем / Schematic representation of the system-level advantages provided by the information technology for developing and validating specifications in the re-engineering of legacy systems

Отже, розроблена інформаційна технологія забезпечує узгоджене поєднання методів і підходів, що дає змогу підвищити ефективність синтезу та перевірки специфікацій вимог до ІТ-проектів у завданнях модернізації складних інформаційних систем.

Методи методики синтезу та валідації специфікацій вимог до ІТ-проектів. Розроблена методологія синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем реалізується на основі сукупності взаємопов'язаних методів, що забезпечують послідовний перехід від аналізу структури системи до формалізації процесів та їх перевірки. Інформаційну технологію побудовано як засіб практичної реалізації цієї методики.

Підсистема реверс-інжинірингу та поведінкового аналізу реалізується двома взаємопов'язаними методами – методом декомпозиції та аналізу legacy-систем і методом синтезу вимог на основі реверс-інжинірингу. Отже, трирівнева підсистемна структура технології деталізується у чотири взаємопов'язані методи, а саме:

- метод декомпозиції та аналізу legacy-систем;

- метод синтезу вимог на основі реверс-інжинірингу;
- метод формалізації бізнес-процесів на основі мереж Петрі;
- метод статистичної валідації специфікацій вимог до ІТ-проектів.

Метод декомпозиції та аналізу legacy-систем. В інформаційній технології удосконалено метод декомпозиції та аналізу legacy-систем, який, на відміну від традиційних підходів до структурного поділу системи, орієнтований не тільки на виділення її наявних функціональних компонентів, а й на підготовку структурованого подання системи з урахуванням цільової логіки подальшої модернізації, синтезу вимог, формалізації бізнес-процесів та статистичної валідації специфікацій вимог до ІТ-проектів [1, 14, 24]. Необхідність застосування такого методу зумовлена тим, що в задачах реінжинірингу legacy-систем великі та багатокомпонентні legacy-системи, зазвичай, характеризуються неповною документацією, прихованими залежностями між модулями та розходженням між задекларованою і фактичною логікою функціонування. За таких умов без попередньої декомпозиції система залишається надто складною для цілісного аналізу, що ускладнює відновлення специфікацій вимог до ІТ-проектів та побудову формальних моделей.

Особливістю удосконаленого методу є те, що декомпозиція виконується не тільки за поточною структурою legacy-системи, а й з урахуванням майбутньої цільової реалізації та архітектури модернізованої системи. Це дає змогу виділяти не тільки наявні функціональні модулі, а й цільові функціональні контури, зони можливої трансформації, майбутні межі підсистем і точки інтеграції, які є важливими для подальшого реінжинірингу legacy-систем. Окрім цього, метод поєднує структурний і поведінковий аналіз, що дає можливість враховувати не тільки склад системи, а й фактичні сценарії її функціонування, критичні залежності та області потенційних функціональних розбіжностей. Це забезпечує перехід від фрагментарних спостережень до структурованого аналітичного подання legacy-системи.

Призначення методу полягає у зменшенні складності аналізу великих і багатокомпонентних legacy-систем шляхом їх впорядкованого поділу на підсистеми, модулі, функціональні блоки та бізнес-процеси, встановлення зв'язків між ними та формування структурованого подання системи як основи для подальшого синтезу вимог і специфікацій вимог до ІТ-проектів.

Вхідними даними методу є результати функціонування legacy-системи, сценарії її роботи, наявна документація, зовнішні реакції системи, відомості про взаємодію користувачів і зовнішніх систем, а також дані про модулі, інтерфейси та потоки оброблення інформації. Результатом застосування методу є декомпозована та структуровано описана модель legacy-системи, що містить виділені функціональні блоки, межі підсистем, ключові бізнес-процеси, залежності між компонентами та критичні точки аналізу, і слугує основою для подальшого синтезу вимог.

Формально результат застосування методу декомпозиції та аналізу legacy-систем можна подати співвідношенням

$$D^* = m_{dec}(L), \quad (2)$$

де: $L = \{l_j, j = \overline{1, n}\}$ – множина вхідних даних для аналізу legacy-системи; m_{dec} – удосконалений метод декомпозиції

та аналізу legacy-систем; D^* – результат декомпозиції та структурованого аналізу legacy-системи.

Основні складові удосконаленого методу охоплюють такі кроки:

- виділення основних функціональних блоків системи;
- визначення меж підсистем і взаємодії між ними;
- ідентифікацію ключових бізнес-процесів;
- встановлення залежностей між компонентами системи;
- виділення цільових функціональних контурів з урахуванням майбутньої архітектури модернізованої системи;
- локалізацію компонентів, що підлягають трансформації, винесенню в окремі сервіси або заміні в процесі модернізації;
- локалізацію критичних точок, у яких можливі функціональні розбіжності або втрата специфікацій вимог до ІТ-проектів;
- формування структурованого опису системи як основи для подальшого поведінкового аналізу.

Вхідними даними для застосування методу є результати функціонування legacy-системи, сценарії її роботи, спостережувані зовнішні реакції, наявна документація, а також відомості про взаємодію користувачів і зовнішніх систем із досліджуванним програмним середовищем. У межах декомпозиції виокремлюються елементи, що мають найбільше значення для реінжинірингу legacy-систем: підсистеми, через які проходять основні потоки даних; модулі, що реалізують критичні бізнес-функції; зовнішні інтерфейси; а також, типові сценарії, в яких проявляється причинно-наслідкова логіка роботи системи. Такий підхід дає змогу перейти від неструктурованої сукупності артефактів legacy-системи до впорядкованого опису її функціональної організації.

Для формального опису методу декомпозиції та аналізу legacy-системи модель системи подано у вигляді кортежу

$$L = \langle X, Y, W, D, \Sigma, R \rangle, \quad (3)$$

де: $X = \{x_j, j = \overline{1, n}\}$ – множина вхідних впливів на систему; $Y = \{y_j, j = \overline{1, n}\}$ – множина вихідних результатів функціонування системи; $W = \{w_j, j = \overline{1, n_W}\}$ – множина спостережуваних операцій; $D = \{d_j, j = \overline{1, n_D}\}$ – множина інформаційних сутностей (даних), що обробляються системою; $\Sigma = \{\varepsilon_j, j = \overline{1, n_\Sigma}\}$ – множина сценаріїв функціонування системи, $R = \{r_j, j = \overline{1, n_R}\}$ – множина залежностей між операціями.

У такому поданні метод дає змогу формалізувати процес переходу від емпіричних даних про функціонування системи до структурованого опису її внутрішньої логіки. Множина X відображає зовнішні впливи, що ініціюють роботу системи, множина Y фіксує результати її функціонування, а множини W, D, Σ та R описують операційний, інформаційний і сценарний рівні аналізу. Саме поєднання цих множин дає змогу подати legacy-систему не тільки як сукупність модулів, а як взаємопов'язану систему процесів, операцій та даних.

Результатом застосування методу є структурована модель legacy-системи, що визначає контекст подальшого аналізу та слугує основою для формування тестових сценаріїв, відновлення функціональних вимог і побудови формальних моделей бізнес-процесів. Отже, метод декомпозиції та аналізу legacy-систем реалізує початковий етап інформаційної технології та забезпечує підготовку вхідних даних для наступного методу – синтезу вимог на основі реверс-інжинірингу.

Метод синтезу вимог на основі реверс-інжинірингу спрямовано на відновлення функціональних специфікацій ІТ-системи на підставі аналізу її поведінки [10, 18, 23]. Метод ґрунтується на застосуванні black-box підходу, у межах якого досліджується реакція системи на контрольовані вхідні впливи.

Синтез вимог містить:

- формування множини тестових сценаріїв;
- виконання серій експериментів над системою;
- фіксацію відповідей між вхідними даними та вихідними результатами;
- узагальнення отриманих сценаріїв у вигляді функціональних залежностей.

Складові методу синтезу вимог подано кортежем

$$M_{req} = \langle S, E, O, A, Q, V \rangle, \quad (4)$$

де: $S = \{s_j, j = \overline{1, n}\}$ – множина функціональних сценаріїв;

$E = \{e_j, j = \overline{1, n}\}$ – множина експериментів із системою;

$O = \{o_j, j = \overline{1, n_o}\}$ – множина спостережень;

$A = \{a_j, j = \overline{1, n_A}\}$ – процедура порівняльного аналізу;

$Q = \{q_j, j = \overline{1, n_Q}\}$ – множина синтезованих вимог;

V – процедура перевірки та уточнення вимог.

Застосування паралельного тестування наявної та модернізованої систем дає змогу виявляти розбіжності у поведінці та уточнювати специфікації [10, 13].

Метод формалізації бізнес-процесів на основі мереж Петрі спрямовано на трансформацію відновлених вимог у формальні моделі бізнес-процесів [9, 11, 17]. Для цього використано апарат мереж Петрі, що забезпечує математично строгий опис динамічної поведінки системи.

Формалізація бізнес-процесів містить:

- відображення станів системи у вигляді позицій;
- подання подій у вигляді переходів;
- моделювання потоків керування;
- опис паралельних і альтернативних сценаріїв виконання;
- врахування механізмів синхронізації.

Формальну модель бізнес-процесу подано мережею Петрі у вигляді кортежу [9, 11, 17]

$$PN = \langle P, T, F, M_0 \rangle, \quad (5)$$

де: $P = \{p_j, j = \overline{1, n}\}$ – множина позицій (станів);

$T = \{t_j, j = \overline{1, n}\}$ – множина переходів (подій);

$F = \{f_j, j = \overline{1, n_F}\}$ – множина дуг, що визначають структуру зв'язків; M_0 – початкове маркування мережі.

Метод дає змогу отримати формальні моделі, придатні для подальшого аналізу та перевірки.

Метод статистичної валідації специфікацій ІТ-проектів спрямовано на перевірку коректності формалізованих специфікацій ІТ-проектів із використанням стохастичного аналізу моделей [2, 4, 12].

Метод містить:

- генерування трас виконання на основі випадкових процесів;
- вибіркового аналізу сценаріїв функціонування;
- оцінювання властивостей моделей без повного перебору станів;
- розрахунок метрик покриття, достовірності та похибки.

Формально стохастичну мережу Петрі можна подати у вигляді кортежу [2, 4, 12]

$$SPN = \langle P, T, F, M_0, \Lambda \rangle, \quad (6)$$

де: P, T, M_0 мають значення, наведені у формулі (5); $F \subseteq (P \times T) \cup (T \times P)$ – множина дуг; $\Lambda: T \rightarrow R_{>0}$ – функція інтенсивностей спрацьовування переходів, яка задає стохастичні параметри моделі.

Застосування методу дає змогу виконувати перевірку великих моделей із зменшеними обчислювальними витратами зі збереженням потрібного рівня достовірності результатів.

Отже, загалом методи інформаційної технології функціонують у послідовному взаємозв'язку:

- 1) Крок 1 – декомпозиція визначає структуру системи.
- 2) Крок 2 – синтез вимог відновлює її поведінку.
- 3) Крок 3 – формалізація створює математичну модель.
- 4) Крок 4 – валідація перевіряє коректність специфікацій ІТ-проектів.

Результати перевірки використовують для уточнення вимог і повторного виконання попередніх кроків, що забезпечує ітеративний характер технології.

Алгоритми реалізації інформаційної технології.

Алгоритми реалізації інформаційної технології розроблено як практичне втілення методики синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем. Розроблену інформаційну технологію реалізовано у вигляді послідовності алгоритмів, що відображають результати попередньо розроблених методів декомпозиції, синтезу вимог, формалізації та статистичної валідації [10, 11, 12]. Сукупність алгоритмів забезпечує покроковий перехід від структурного аналізу системи до формування специфікацій ІТ-проектів їх формалізації та статистичної перевірки.

До складу інформаційної технології входять:

- алгоритм декомпозиції та аналізу legacy-системи;
- алгоритм синтезу вимог на основі реверс-інжинірингу;
- алгоритм формалізації бізнес-процесів на основі мереж Петрі;
- алгоритм статистичної валідації специфікацій ІТ-проектів.

Алгоритм декомпозиції та аналізу реалізує початковий етап інформаційної технології та спрямований на виділення функціональної структури legacy-системи.

Алгоритм містить такі кроки:

- 1) Крок 1 – виділення меж досліджуваної системи та її зовнішнього оточення.
- 2) Крок 2 – визначення основних функціональних підсистем і модулів.
- 3) Крок 3 – встановлення зв'язків між підсистемами, потоками даних і керівними впливами.
- 4) Крок 4 – ідентифікацію ключових бізнес-процесів, що підлягають реінжинірингу.
- 5) Крок 5 – локалізацію критичних точок, у яких можливі функціональні розбіжності або втрата специфікацій ІТ-проектів.
- 6) Крок 6 – формування структурованого опису системи як основи для подальшого поведінкового аналізу.

Результатом виконання алгоритму є декомпонована модель legacy-системи, що задає контекст для синтезу вимог і побудови тестових сценаріїв.

Алгоритм синтезу вимог реалізує відновлення специфікацій ІТ-проектів на основі аналізу поведінки системи та ґрунтується на black-box підході [10, 13].

Алгоритм містить такі кроки:

- 1) Крок 1 – формування множини вхідних сценаріїв, що охоплюють основні режими функціонування системи.
- 2) Крок 2 – запуск експериментів над наявною системою та фіксацію її реакцій.
- 3) Крок 3 – накопичення пар "вхідний вплив – вихідний результат".
- 4) Крок 4 – групування однорідних реакцій системи та виділення повторюваних поведінкових шаблонів.
- 5) Крок 5 – побудову узагальнених функціональних залежностей між вхідними даними, станами процесів та результатами.

6) Крок 6 – виконання паралельного тестування наявної та модернізованої систем.

7) Крок 7 – виявлення розбіжностей у поведінці та документування невідповідностей.

8) Крок 8 – уточнення функціональних вимог і формування узгодженого набору специфікацій ІТ-проектів.

Результатом виконання алгоритму є синтезована множина вимог і сценаріїв виконання, що відображають реальну поведінку legacy-системи.

Алгоритм формалізації реалізує перехід від відновлених сценаріїв до формального опису процесів [11, 17]. Алгоритм містить такі кроки:

- 1) Крок 1 – виділення станів бізнес-процесу на основі синтезованих вимог.
- 2) Крок 2 – відображення станів у вигляді позицій мережі Петрі.
- 3) Крок 3 – виділення подій, операцій і умов переходу між станами.
- 4) Крок 4 – подання подій у вигляді переходів.
- 5) Крок 5 – побудову дуг між позиціями та переходами відповідно до логіки виконання процесу.
- 6) Крок 6 – відображення альтернативних гілок, паралельних підпроцесів і механізмів синхронізації.
- 7) Крок 7 – перевірку структурної коректності побудованої моделі.
- 8) Крок 8 – формування завершеної формальної моделі бізнес-процесу.

Результатом виконання алгоритму є формальна модель процесу на основі мережі Петрі, придатна для подальшої перевірки властивостей специфікації.

Алгоритм статистичної валідації реалізує перевірку коректності специфікацій ІТ-проектів на підставі стохастичного аналізу побудованих моделей [2, 4, 12].

Алгоритм містить такі кроки:

- 1) Крок 1 – задавання початкових параметрів перевірки: кількості ітерацій, критеріїв зупинки та порогових значень метрик.
- 2) Крок 2 – генерування множини трас виконання на основі стохастичних переходів у моделі.
- 3) Крок 3 – фіксацію послідовностей спрацьовування переходів і досягнутих станів.
- 4) Крок 4 – зіставлення згенерованих трас із очікуваними сценаріями специфікації.
- 5) Крок 5 – визначення випадків невідповідності, конфліктів, пропусків або надлишкових гілок.
- 6) Крок 6 – розрахунок метрик покриття, достовірності та похибки.
- 7) Крок 7 – оцінювання результатів перевірки відповідно до заданих критеріїв.
- 8) Крок 8 – прийняття рішення щодо коректності специфікації або потреби її уточнення.

Результатом виконання алгоритму є кількісно обґрунтований висновок про якість специфікації ІТ-проектів та перелік виявлених невідповідностей.

Принциповою особливістю є те, що вказані алгоритми функціонують не ізольовано, а як послідовні компоненти єдиного інтегрованого алгоритму реалізації інформаційної технології синтезу та валідації специфікацій вимог до ІТ-проектів.

Інтегрований алгоритм передбачає виконання таких кроків:

- 1) Крок 1 – декомпозиція та структурний аналіз legacy-системи.
- 2) Крок 2 – побудова тестових сценаріїв для дослідження її поведінки.
- 3) Крок 3 – синтез вимог на основі результатів black-box аналізу та паралельного тестування.
- 4) Крок 4 – трансформація відновлених вимог у формальні моделі бізнес-процесів.

5) Крок 5 – статистична валідація специфікацій ІТ-проектів за згенерованими трасами виконання.

6) Крок 6 – аналіз результатів перевірки.

7) Крок 7 – уточнення вимог і повторне виконання попередніх етапів у разі виявлення невідповідностей.

Розроблені алгоритми мають такі особливості:

- модульність, що дає змогу виконувати окремі етапи незалежно;
- ітеративність, що забезпечує повторне уточнення специфікацій ІТ-проектів;
- формалізованість, що створює основу для автоматизації;
- масштабованість, що дає змогу застосовувати алгоритми до складних legacy-систем.

Вказані особливості реалізації алгоритмів забезпечили придатність інформаційної технології до практичного використання у задачах реінжинірингу та модернізації ІТ-проектів. Отже, інформаційна технологія передбачає реалізацію ітеративного процесу, у межах якого відбувається послідовне підвищення повноти та узгодженості специфікацій вимог до ІТ-проектів.

Програмна реалізація. Програмну реалізацію розробленої інформаційної технології виконано як інструментальне втілення методики синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем у вигляді експериментального кросплатформеного програмного комплексу, призначеного для підтримки процесів аналізу legacy-систем, формалізації бізнес-процесів та статистичної валідації специфікацій ІТ-проектів. Реалізація орієнтована не на створення завершеного промислового продукту, а на практичне підтвердження працездатності розроблених методів, алгоритмів та інформаційної технології загалом [10, 11, 12].

Архітектуру програмного комплексу побудовано відповідно до структури інформаційної технології. Окремі функціональні модулі відповідають крокам декомпозиції та аналізу legacy-системи, синтезу вимог на основі реверс-інжинірингу, формалізації бізнес-процесів на основі мереж Петрі та статистичної валідації специфікацій ІТ-проектів. Такий підхід забезпечив узгодженість між методичною та програмною складовими дослідження і дав змогу реалізувати наскрізний цикл оброблення даних – від фіксації поведінкових сценаріїв до отримання кількісної оцінки коректності специфікації.

Програмний комплекс реалізовано засобами Python, що дало змогу поєднати оброблення вхідних даних, побудову внутрішніх структур моделі, генерування трас виконання та статистичний аналіз результатів перевірки в єдиному середовищі [12]. Вибір цього інструментального середовища зумовлено його придатністю до швидкого розроблення дослідних програмних засобів, підтримкою модульної архітектури та наявністю бібліотек для математичного і статистичного опрацювання даних. Для роботи із вхідними наборами сценаріїв та результатами перевірки застосовано табличні та структуровані формати подання даних, що спростило автоматизацію експериментальних досліджень.

На етапі декомпозиції та аналізу legacy-систем програмний комплекс забезпечує введення і впорядкування інформації в базу знань про функціональні компоненти legacy-системи, її підсистеми, сценарії виконання та ключові точки взаємодії. Ці знання використовують як основу для формування тестових сценаріїв і подальшого синтезу вимог. Підсистема синтезу вимог реалізує оброблення результатів black-box аналізу та паралель-

ного тестування, зіставлення реакцій наявної та модернізованої системи та фіксацію виявлених розбіжностей. У межах цієї підсистеми виконують узагальнення поведінкових шаблонів і формують структурований опис вимог, що далі передається на етап формалізації.

Підсистема формалізації виконує побудову внутрішнього подання бізнес-процесів у термінах мереж Петрі. Реалізація цієї частини комплексу забезпечує формування множини позицій, переходів і дуг, відображення паралельних та альтернативних сценаріїв, а також побудову зв'язків між елементами моделі відповідно до логіки процесу. Унаслідок формується модель, придатна для подальшого аналізу та статистичної перевірки. Така програмна реалізація дала змогу відтворити поведінку досліджуваних процесів у строгому формальному вигляді та усунути неоднозначності, притаманні текстовому опису специфікацій ІТ-проектів.

Підсистема статистичної валідації реалізує генерування трас виконання, накопичення результатів моделювання та обчислення кількісних характеристик перевірки. У програмному комплексі реалізовано механізм багаторазового проходження моделі з фіксацією послідовностей спрацьовування переходів і досягнутих станів. Це дало змогу перейти від повного перебору станів до вибіркового аналізу поведінки моделі. На підставі згенерованих трас виконують розрахунок метрик покриття, достовірності та похибки, після чого формують висновок про коректність специфікації або необхідність її уточнення. Вхідними даними програмного комплексу є сценарії поведінки системи, результати тестування, журнали подій та структуровані описи вимог, а вихідними – формалізовані моделі, згенеровані траси виконання, значення метрик C , P , E та висновки щодо коректності специфікації.

Результати роботи програмного комплексу подаються у вигляді структурованих даних, що містять згенеровані сценарії, виявлені невідповідності, кількісні показники валідації та підсумкову оцінку специфікації. Це дало змогу використовувати програмну реалізацію не тільки як засіб перевірки окремих моделей, а і як інструмент підтримки прийняття рішень у процесі реінжинірингу ІТ-проектів.

Аналіз отриманих результатів дослідження. Експериментальне дослідження виконано задля перевірки працездатності розробленої інформаційної технології та оцінювання ефективності її основних складових у задачах реінжинірингу legacy-систем. Дослідження охопило повний цикл застосування технології – від декомпозиції та аналізу системи до статистичної валідації формалізованих специфікацій ІТ-проектів. Основну увагу зосереджено на тому, наскільки коректно відновлені вимоги відображають реальну поведінку системи, наскільки адекватно побудовані формальні моделі описують бізнес-процеси та якою мірою статистична перевірка знижує обчислювальні витрати порівняно з повною формальною верифікацією.

Експериментальні результати оцінено на моделях бізнес-процесів різної складності, для яких послідовно виконано декомпозицію, синтез вимог, формалізацію на основі мереж Петрі та статистичну валідацію. Порівняння проводили між повною формальною перевіркою моделей і вибірковою статистичною перевіркою за згенерованими трасами виконання. Для експериментального оцінювання було сформовано три моделі різної

складності на основі побудованої мережі Петрі CCR. Структурні характеристики моделей наведено в табл. 1, де мірою складності обрано кількість позицій, переходів і дуг.

Табл. 1. Характеристики експериментальних моделей виконання Consumer Credit Report на основі мережі Петрі / Consumer Credit Report processing experimental the Petri net model characteristic

Мо- дель	Назва	К-сть позицій	К-сть переходів	К-сть дуг	L_{max}
M1	Базова модель валідації запиту	28	36	86	55
M2	Ядро процесу формування базового CCR	58	74	182	170
M3	Повна production-like модель CCR	128	168	412	520

Модель M1 охоплює тільки базові етапи валідації запиту, модель M2 відображає ядро процесу формування базового CCR для одного заявника, а модель M3 представляє повну модель CCR, наближену до виробничого сценарію, із підтримкою joint-запитів, ancillary-сервісів, inquiry posting та billing. Для кожної моделі виконано bounded exhaustive traversal із обмеженням L_{max} для побудови еталонної множини трас Ω_{full} , а також статистичну генерування трас $\Omega_{stat}(n)$, методом Монте-Карло для різних значень n . На цій основі обчислено показники покриття, частку успішних та помилкових трас, середню довжину траси, похибку, достовірність і обчислювальну вартість статистичної перевірки.

На першому кроці експериментів оцінено результати декомпозиції та поведінкового аналізу legacy-систем. Отримані дані показали, що структуризація системи за функціональними підсистемами та бізнес-процесами спростила локалізацію критичних сценаріїв і дала змогу зосередити подальший аналіз на найбільш значущих ділянках функціонування. Застосування black-box підходу в поєднанні з паралельним тестуванням наявної та модернізованої систем забезпечило виявлення розбіжностей у поведінці, які не були очевидними з наявної документації [10, 13].

У табл. 2 наведено статистичні дані щодо кількості випадків у вибірці, числа збігів, доданих, видалених і змінених елементів, а також їх відносних часток для кожної ітерації паралельного тестування.

Унаслідок уточнено множини функціональних вимог і сформовано узгоджені сценарії виконання, що стали основою для подальшої формалізації.

На другому кроці досліджено адекватність формалізації бізнес-процесів на основі мереж Петрі [9, 11, 17]. Побудовані моделі відобразили послідовні, паралельні та альтернативні гілки виконання процесів, а також механізми синхронізації та переходів між станами (рис. 3).

Порівняння логіки моделей із поведінковими сценаріями, отриманими на попередньому кроці, підтвердило, що формалізація забезпечила цілісне й несуперечливе подання процесів. Це дало змогу перейти від текстового опису вимог до строгого формального подання, придатного для подальшої валідації.

На третьому кроці виконано статистичну валідацію специфікацій ІТ-проектів на основі згенерованих трас виконання [2, 4, 12]. Для цього використовувалася імітаційна генерація трас $\tau = \{\tau_1, \tau_2, \dots\}$, де кожна траса відповідає допустимій послідовності переходів мережі. Оцінювання виконано за метриками покриття C , достовірності P та похибки E , а результати наведено у табл. 3.

Табл. 2. Кількісні результати ітеративного паралельного тестування на різних етапах узгодження специфікації /
Quantitative results of iterative parallel testing at different stages of specification alignment

Ітерація	Дата	Кількість випадків, шт.	Збіги	Додані правила	Видалені правила	Змінені правила	Збіги, %	Додані, %	Видалені, %	Змінені, %
Перша	15.03.23	35038	249	6398	0	28391	0,71	18,26	0,00	81,03
	16.03.23	37172	270	6469	0	30433	0,73	17,40	0,00	81,87
	17.03.23	39887	305	6377	0	33205	0,76	15,99	0,00	83,25
	20.03.23	35674	223	6681	0	28770	0,63	18,73	0,00	80,65
Друга	23.03.23	30450	249	2091	0	28110	0,82	6,87	0,00	92,32
	24.03.23	29180	217	916	397	27650	0,74	3,14	1,36	94,76
Третя	27.03.23	36485	223	780	288	35194	0,61	2,14	0,79	96,46
	28.03.23	16323	206	645	87	15385	1,26	3,95	0,53	94,25
	30.03.23	23680	182	775	326	22397	0,77	3,27	1,38	94,58
	04.04.23	35514	16547	0	294	18673	46,59	0,00	0,83	52,58
Четверта	11.04.23	31392	15386	0	215	15791	49,01	0,00	0,68	50,30
	18.04.23	33220	15619	1	262	17338	47,02	0,00	0,79	52,19
	03.05.23	32817	32716	0	18	83	99,69	0,00	0,05	0,25
П'ята	04.05.23	34305	34172	0	21	112	99,61	0,00	0,06	0,33

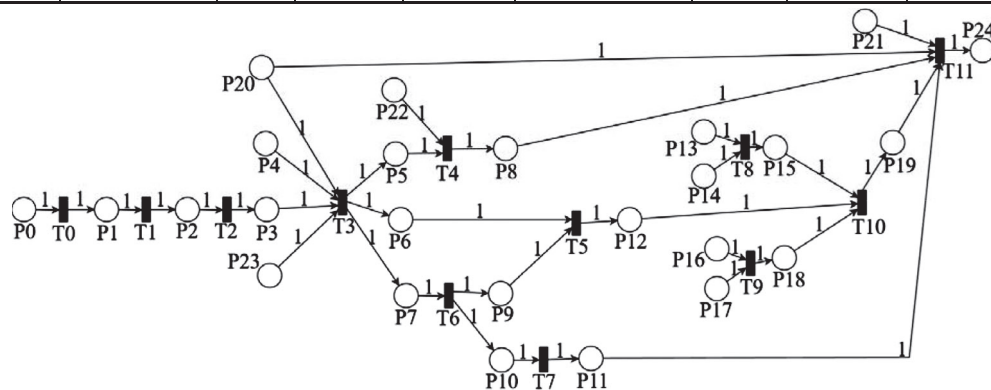


Рис. 3. Фрагмент моделі процесу генерування Consumer Credit Report на основі мережі Петрі / Fragment of the Consumer Credit Report generation process model based on a Petri net [11]

Табл. 3. Результати порівняння повної перевірки та статистичної валідації моделі мережі Петрі методом Монте-Карло / Comparison of full verification and Monte Carlo – based statistical validation results for the Petri net model

	Режим	Перевірені переходи	Покриття, С	Достовірність, Р	Похибка, Е	Обчислювальна вартість
Модель М1	Повна перевірка	36	1,000000	1,000000	0,000000	9225
	Монте-Карло, N=100	29	0,805556	0,960000	0,040000	1948
	Монте-Карло, N=250	32	0,888889	0,964000	0,036000	4784
	Монте-Карло, N=500	35	0,972222	0,998000	0,002000	9585
	Монте-Карло, N=1000	34	0,944444	0,998000	0,002000	19164
	Монте-Карло, N=2000	36	1,000000	0,995000	0,005000	38348
	Монте-Карло, N=5000	36	1,000000	0,991600	0,008400	96508
Модель М2	Повна перевірка	74	1,000000	1,000000	0,000000	30266
	Монте-Карло, N=100	59	0,797297	0,990000	0,010000	3091
	Монте-Карло, N=250	71	0,959459	0,994000	0,006000	7500
	Монте-Карло, N=500	69	0,932432	1,000000	0,000000	14975
	Монте-Карло, N=1000	72	0,972973	0,998000	0,002000	30057
	Монте-Карло, N=2000	74	1,000000	0,998500	0,001500	60217
	Монте-Карло, N=5000	74	1,000000	0,997800	0,002200	150898
Модель М3	Повна перевірка	168	1,000000	1,000000	0,000000	82396
	Монте-Карло, N=100	148	0,880952	0,991305	0,008695	5444
	Монте-Карло, N=250	148	0,880952	0,990695	0,009305	13525
	Монте-Карло, N=500	148	0,880952	0,987305	0,012695	27245
	Монте-Карло, N=1000	155	0,922619	0,998305	0,001695	55253
	Монте-Карло, N=2000	157	0,934524	0,999805	0,000195	109471
	Монте-Карло, N=5000	157	0,934524	0,996305	0,003695	273265
	Монте-Карло, N=10000	157	0,934524	0,998705	0,001295	548231
	Монте-Карло, N=20000	157	0,934524	0,999495	0,000505	1098493

Результати етапів проведеного експерименту показали, що зі збільшенням кількості згенерованих трас досягається стабілізація значення покриття та зменшення похибки, що підтверджує ефективність статистичного підходу. У проведених етапах дослідження було зафік-

совано зменшення обчислювальних витрат у межах приблизно 1,5-3,0 рази залежно від складності моделі та параметрів генерування трас, що узгоджується з результатами, наведеними в попередніх працях [12]:

1) Модель М1: статистична валідація не дає виграву, бо модель занадто мала і повний перебір дешевший.

- 2) Модель M2: виграш є для малих і середніх N , далі він зникає.
- 3) Модель M3: виграш найвиразніший, бо модель складніша, але за дуже великого N статистична перевірка теж стає дорожчою за baseline.

Отримані результати підтвердили не тільки працездатність окремих методів, а й ефективність їх інтеграції в межах єдиної інформаційної технології. Узгоджене поєднання декомпозиції системи, синтезу вимог, формалізації бізнес-процесів і статистичної валідації забезпечило послідовне уточнення специфікацій ІТ-проектів та підвищення їх узгодженості з реальною поведінкою системи. Важливо, що результати експериментів засвідчили практичну придатність технології для задач реінжинірингу ІТ-проектів, у яких повна формальна перевірка є надто ресурсомісткою або практично неможливою.

Отже, експериментальне дослідження підтвердило, що розроблена інформаційна технологія забезпечує цілісний і масштабований підхід до синтезу та валідації специфікацій вимог до ІТ-проектів у процесі модернізації legacy-систем. Її застосування дає змогу підвищити повноту та узгодженість специфікацій вимог до ІТ-проектів формалізувати логіку процесів і виконати їх перевірку з прийнятними обчислювальними витратами.

Отримані результати дослідження створюють науково-практичну основу для подальшого розвитку засобів синтезу та валідації специфікацій вимог до ІТ-проектів і можуть бути використані під час модернізації корпоративних, сервісних та інших складних legacy-систем.

Обговорення результатів дослідження. У дослідженні [15] автори розглянули використання model-driven engineering для автоматизованого переходу від програмного коду до специфікацій у формах UML та OCL, що підтверджує доцільність використання модельно-орієнтованих підходів для подальшої формалізації програмних систем. На відміну від цього підходу, у нашому дослідженні формалізацію виконано не як перехід між мовами програмування, а як перехід від поведінкового аналізу legacy-системи до специфікації, описаної через бізнес-процеси та мережі Петрі, що дало змогу безпосередньо інтегрувати формалізацію зі статистичною валідацією специфікацій ІТ-проектів.

У роботі [16] автори дослідили стохастичні мережі Петрі з silent transitions і показали можливість аналітичного аналізу трас та ймовірностей у process mining задачах. Результати, які отримали вони, є близькими до нашого підходу в частині стохастичного аналізу поведінки моделі, однак у нашому дослідженні статистичну валідацію спрямовано не на загальний стохастичний аналіз процесів, а на перевірку коректності специфікацій, синтезованих у задачах реінжинірингу legacy-систем, що розширює сферу застосування Petri net-підходу.

У роботі [19] автори запропонували data-driven framework для оцінювання бізнес-процесів у комерційних банках із використанням process mining, соціально-мережевого аналізу та багатокритеріального оцінювання. Ці результати підтверджують зростання ролі даних і журналів подій у дослідженні реальних процесів. На відміну від такого підходу, у нашому дослідженні увагу зосереджено не тільки на оцінюванні процесів, а й на відновленні вимог, їх формалізації та подальшій статистичній перевірці, що дає змогу перейти від аналізу наявних процесів до побудови валідованої специфікації нової системи.

У дослідженні [21] автори розглянули можливості великих мовних моделей для semantics-aware process mining tasks, тобто для завдань, у яких важливим є не тільки порядок дій, а й змістовий зв'язок між ними. Цей підхід підтверджує актуальність інтелектуального аналізу процесів на основі їх семантики. Однак у нашому дослідженні зроблено акцент на формальному поданні процесів і статистичній валідації специфікацій ІТ-проектів, що забезпечує не тільки семантичне узгодження, а й можливість кількісної перевірки коректності бізнес-логіки.

У роботі [20] автори запропонували object-centric підхід до операційного моніторингу процесів, який орієнтований на виявлення проблем у складних багатосутнісних бізнес-процесах. Отримані результати демонструють важливість сучасних методів процесного моніторингу для аналізу поведінки реальних систем. На відміну від цього, у нашому дослідженні процесний аналіз інтегровано з декомпозицією legacy-системи, синтезом вимог, формалізацією бізнес-процесів і валідацією специфікацій ІТ-проектів, що дає змогу використовувати результати моніторингу не тільки для виявлення проблем, а і для побудови нової формалізованої специфікації.

У роботі [22] автори дослідили business process re-engineering як засіб оптимізації діяльності органів місцевого самоврядування та показали, що реінжиніринг процесів може істотно покращувати організаційну ефективність. Ці результати підтверджують загальну доцільність процесного реінжинірингу legacy-систем. Однак у нашому дослідженні, на відміну від організаційно-управлінського фокусу цієї роботи, основну увагу приділено формалізації та валідації специфікацій ІТ-проектів, що дало змогу поєднати аналіз процесів із побудовою математичних моделей і засобами їх перевірки.

Отже, внаслідок обговорення результатів дослідження підтверджено доцільність проведення саме цього дослідження, оскільки наявні в сучасній науковій літературі підходи переважно розглядають окремі аспекти реінжинірингу legacy-систем, зокрема – модельно-орієнтований реверс-інжиніринг, process mining, process monitoring, stochastic Petri nets або business process re-engineering, тоді як у нашому дослідженні розроблено цілісну інформаційну технологію, яка поєднує декомпозицію legacy-систем, синтез вимог, формалізацію бізнес-процесів на основі мереж Петрі та статистичну валідацію специфікацій ІТ-проектів у межах єдиного ітеративного циклу, що дало змогу отримати результати, які не зводяться до окремого аналізу процесів, а забезпечують узгоджене розроблення та перевірку специфікацій вимог до ІТ-проектів.

Отже, внаслідок виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

Наукова новизна отриманих результатів дослідження – вперше розроблено методику синтезу та валідації специфікацій вимог до ІТ-проектів у задачах реінжинірингу legacy-систем, яка інтегрує метод декомпозиції та аналізу legacy-систем, метод синтезу вимог на основі реверс-інжинірингу, метод формалізації бізнес-процесів на основі мереж Петрі та метод статистичної валідації специфікацій ІТ-проектів у межах єдиного ітеративного циклу, що дало змогу підвищити повноту й узгодженість специфікацій та зменшити обчислювальні витрати їх перевірки.

Практична значущість результатів дослідження – запропоновану методику можна використати в ІТ-проєктах для автоматизації процесів синтезу, уточнення та валідації специфікацій складних ІТ-систем, зокрема – під час реінжинірингу legacy-систем, модернізації бізнес-процесів і контролю узгодженості вимог, а також для зменшення обчислювальних витрат перевірки специфікацій ІТ-проєктів за рахунок застосування статистичної валідації.

Висновки / Conclusions

Розроблено інформаційну технологію синтезу та валідації специфікацій вимог до ІТ-проєктів у задачах реінжинірингу legacy-систем, яка дає змогу поєднати поведінковий аналіз, формальне моделювання та масштабовану статистичну їх перевірку. За результатами проведеного дослідження можна зробити такі основні висновки.

1. Проаналізовано обмеження традиційних підходів до синтезу та валідації специфікацій вимог до ІТ-проєктів у задачах реінжинірингу legacy-систем, зокрема – їх фрагментарність, недостатню орієнтацію на поведінковий аналіз системи, відсутність цілісного поєднання реверс-інжинірингу, формального моделювання та масштабованої перевірки специфікацій ІТ-проєктів.
2. Проаналізовано та систематизовано наявні й розроблені в межах дослідження методи та алгоритми, що забезпечують синтез і валідацію специфікацій вимог до ІТ-проєктів у задачах реінжинірингу legacy-систем, а саме: метод декомпозиції та аналізу legacy-систем, метод синтезу вимог на основі реверс-інжинірингу, метод формалізації бізнес-процесів на основі мереж Петрі, метод статистичної валідації специфікацій ІТ-проєктів, а також алгоритми їх реалізації.
3. Удосконалено метод декомпозиції та аналізу legacy-систем, а також метод синтезу вимог на основі реверс-інжинірингу шляхом поєднання поведінкового аналізу за принципом black-box, паралельного тестування та структуризації функціональних залежностей, що дало змогу підвищити повноту й узгодженість специфікацій ІТ-проєктів, відновлених у процесі реінжинірингу legacy-систем.
4. Розроблено методику синтезу та статистичної валідації специфікацій ІТ-проєктів у задачах реінжинірингу legacy-систем, яка об'єднує методи декомпозиції та аналізу системи, синтезу вимог, формалізації бізнес-процесів на основі мереж Петрі, а також алгоритми їх реалізації, що дало змогу забезпечити цілісний ітеративний процес розроблення, уточнення та перевірки синтезованих специфікацій ІТ-проєктів.
5. Розроблено інформаційну технологію практичної реалізації методики синтезу та валідації специфікацій вимог до ІТ-проєктів у задачах реінжинірингу legacy-систем, яка об'єднує відповідні методи, алгоритми та програмні засоби і забезпечує практичну підтримку цілісного ітеративного процесу аналізу, формалізації, перевірки та уточнення специфікацій вимог до ІТ-проєктів.

References

1. Abdellatif, M., Shatnawi, A., Mili, H., Moha, N., El Boussaidi, G., Hecht, G., Privat, J., & Guéhéneuc, Y.-G. (2021). *A taxonomy of service identification approaches for legacy software systems modernization*. *Journal of Systems and Software*, 173, article ID 110868. <https://doi.org/10.1016/j.jss.2020.110868>
2. Budde, C. E., Hartmanns, A., Meggendorfer, T., Weininger, M., & Wienhöft, P. (2024). *Sound statistical model checking for probabilities and expected rewards*. arXiv. <https://doi.org/10.48550/arXiv.2411.00559>
3. Cazzola, W., & Favalli, L. (2024). Software modernization powered by dynamic language product lines. *Journal of Systems and*

- Software*, 218, article ID 112188. <https://doi.org/10.1016/j.jss.2024.112188>
4. Dubois, T., Larsen, K. G., & Srba, J. (2025). *Statistical model checking of stochastic timed-arc Petri nets*. In *Application and Theory of Petri Nets and Concurrency* (pp. 174–196). Springer. https://doi.org/10.1007/978-3-031-94634-9_9
5. Fávero, L. F., de Almeida, N. R., & Affonso, F. J. (2025). A systematic mapping study on the modernization of legacy systems to microservice architecture. *Applied System Innovation*, 8(4), article number 86. <https://doi.org/10.3390/asi8040086>
6. Guez Assuncao, W. K., Marchezan, L., Arkoh, L., Egyed, A., & Ramler, R. (2025). Contemporary software modernization: Strategies, driving forces, and research opportunities. *ACM Transactions on Software Engineering and Methodology*, 34(5), article ID 142, 1–35. <https://doi.org/10.1145/3708527>
7. Hasan, M. H., Osman, M. H., Admodisastro, N. I., & Muhammad, M. S. (2023). *Legacy systems to cloud migration: A review from the architectural perspective*. *Journal of Systems and Software*, 202, article ID 111702. <https://doi.org/10.1016/j.jss.2023.111702>
8. Irani, Z., Abril, R. M., Weerakkody, V., Omar, A., & Sivarajah, U. (2023). The impact of legacy systems on digital transformation in European public administration: Lesson learned from a multi case analysis. *Government Information Quarterly*, 40(1), article ID 101784. <https://doi.org/10.1016/j.giq.2022.101784>
9. Kamal, A., Lubis, M., Dewi, F., & Constantino, M. G. (2023). A verification and validation of business process modeling using Petri nets (Case study of Telkom University FRI Laboratory Business Architecture Domain). *Electronic Integrated Computer Algorithm Journal*, 1(1), 26–31. <https://doi.org/10.62123/enigma.v1i1.5>
10. Kernysky, O. B., & Teslyuk, V. M. (2023). The synthesis method for specifications and requirements in the process of it project reengineering. *Ukrainian Journal of Information Technology*, 5(2), 01–08. <https://doi.org/10.23939/ujit2023.02.001>
11. Kernysky, O., & Kernysky, A. (2025). A Petri net-based approach to extracting and validating requirements for modelling and modernizing archaic IT systems. *Computer. Dysign Systems. Theory and practice*, 7(2), 132–143. <https://doi.org/10.23939/cds2025.02.132>
12. Kernysky, O., & Teslyuk, V. (2026). Modeling Specification Validation in IT Project Reengineering Based on Statistical Analysis of Petri Nets. *Scientific Bulletin of UNFU*, 36(1), 96–104. <https://doi.org/10.36930/40360111>
13. Kernysky, O., Kernysky, A., Melnyk, M., Łukaszewicz, A., Pytel, K., & Banas, M. (2024). Applying software development black-box, grey-box and white box reverse engineering frameworks to the mechanical industry, CAD in Machinery Design. *Oficina Wydawnicza Politechniki Białostockiej, Białystok* 2024, 233–240. <https://doi.org/10.24427/978-83-68077-19-3>
14. Kirschner, Y. R., Gstür, M., Sağlam, T., Weber, S., & Koziol, A. (2025). *Retriever: A view-based approach to reverse engineering software architecture models*. *Journal of Systems and Software*, 220, article ID 112277. <https://doi.org/10.1016/j.jss.2024.112277>
15. Lano, K., & Siala, H. (2024). Using model-driven engineering to automate software language translation. *Automated Software Engineering*, 31(1), article number 20. <https://doi.org/10.1007/s10515-024-00419-y>
16. Leemans, S. J. J., Maggi, F. M., & Montali, M. (2024). Enjoy the silence: Analysis of stochastic Petri nets with silent transitions. *Information Systems*, 124, article ID 102383. <https://doi.org/10.1016/j.is.2024.102383>
17. Medina-Garcia, S., Medina-Marin, J., Montañó-Arango, O., Gonzalez-Hernandez, M., & Hernandez-Gress, E. S. (2023). *A Petri net approach for business process modeling and simulation*. *Applied Sciences*, 13(20), article ID 11192. <https://doi.org/10.3390/app132011192>
18. Miranda, N. A., & Ergin, N. (2025). A reverse-MBSE approach to legacy systems: CODIS as a case study. *Procedia Computer Science*, 268, 139–147. <https://doi.org/10.1016/j.procs.2025.08.190>

19. Ni, Z., Xiao, B., & Li, Y. (2025). Data-Driven Business Process Evaluation in Commercial Banks: Multi-Dimensional Framework with Hybrid Analytical Approaches. *Systems*, 13(4), article number 256. <https://doi.org/10.3390/systems13040256>
20. Park, G., & van der Aalst, W. M. P. (2025). Operational process monitoring: An object-centric approach. *Computers in Industry*, 164, article ID 104170. <https://doi.org/10.1016/j.com-pind.2024.104170>
21. Rebmann, A., Schmidt, F., Glavaš, G., & van der Aa, H. (2025). On the potential of large language models to solve semantics-aware process mining tasks. *Process Science*, 2, article number 10. <https://doi.org/10.1007/s44311-025-00019-3>
22. Sayuti, M., Syairudin, B., & Gunarta, I. K. (2025). Enhancement of business processes through re-engineering to optimize the performance of local government in Central Sulawesi province. *Cogent Social Sciences*, 11(1), article ID 2542922. <https://doi.org/10.1080/23311886.2025.2542922>
23. Siala, H. A., Lano, K., & Alfraihi, H. (2024). Model-driven approaches for reverse engineering: A systematic literature review. *IEEE Access*, 12, 62558–62580. <https://doi.org/10.1109/AC-CESS.2024.3394732>
24. Terry, S., & Chandrasekar, V. (2025). Systems engineering barriers to legacy system evolution: Legacy system assessment. *Systems Engineering*, 28(2), 207–223. <https://doi.org/10.1002/sys.21788>

O. B. Kernytskyy, Yu. V. Opoťyak, V. M. Teslyuk

Lviv Polytechnic National University, Lviv, Ukraine

INFORMATION TECHNOLOGY OF SPECIFICATIONS SYNTHESIS AND VALIDATION IN THE PROCESS OF IT PROJECT REENGINEERING

Legacy system reengineering requires more than software renewal; it also requires the recovery, refinement, and validation of IT project specifications that accurately reflect actual system behavior. The purpose of this study is to develop an information technology for the synthesis and validation of IT project specifications in legacy system reengineering tasks. The study addresses this problem by integrating legacy system decomposition and analysis, requirement synthesis based on reverse engineering, business process formalization with Petri nets, and statistical IT project specification validation within a single iterative cycle. Moreover, the study develops a methodology for the synthesis and validation of IT project specifications, defines its methodological principles, establishes the structure of the information technology for its practical implementation, develops a coherent set of interrelated methods and algorithms, and provides an experimental cross-platform software implementation. The main results show that the proposed information technology supports the complete transition from behavioral analysis of a legacy system to formalized and validated specifications. In particular, the study introduces tuple-based mathematical models for legacy system decomposition and analysis, requirement synthesis, business process formalization, and statistical IT project specification validation. The benefits of this study are demonstrated through the Consumer Credit Report specification case (SRS v1.1.37), where the technology was applied to validate input requests, business logic, and output report generation. The study derived requirements and validation rules from reverse engineering results, validated them through parallel testing, formalized them as a Petri net, and then checked them by stochastic simulation of system behavior. As a result, the experiments confirmed the operability of the developed information technology and showed that statistical validation reduces computational costs by approximately 1.5 to 3 times compared with full formal verification while preserving the required level of result reliability. Consequently, the developed information technology can support legacy system modernization by improving specification consistency, strengthening validation procedures, and reducing the cost of formal checking in complex corporate IT environments.

Keywords: reverse engineering; requirements synthesis; analysis of IT project specifications; parallel testing; black-box model; requirements traceability matrix; functional discrepancy analysis; Petri nets; statistical validation of IT project specifications.