МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧОРНОМОРСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ПЕТРА МОГИЛИ

Казакевич Дмитро Андрійович

004.052.2

СТАТИСТИЧНИЙ АНАЛІЗ ВИБІРКИ НАЙБІЛЬШ ПОШИРЕНИХ ПОМИЛОК В ІТ-ПРОЕКТАХ ДЛЯ ОПТИМІЗАЦІЇ РОБОТИ КОМАНД РОЗРОБНИКІВ

122 – Комп’ютерні науки

Автореферат
магістерської кваліфікаційної роботи на здобуття освітньої кваліфікації
«Магістр комп’ютерних наук»

Миколаїв – 2021

Магістерська кваліфікаційна робота є рукопис.

Робота виконана в Чорноморському національному університеті імені Петра Могили Міністерства освіти і науки України на кафедрі інтелектуальних інформаційних систем.

Науковий керівник: к.ф.-м.н., доцент (б.в.з.) кафедри інтелектуальних інформаційних систем Кулаковська Інесса Василівна

Рецензент: д.т.н., професор кафедри комп'ютерної інженерії Кутковецький Валентин Якович

Захист відбудеться «23» лютого 2021 р. о 9^{00} год. на засіданні екзаменаційної комісії (ауд. 2-403) у Чорноморському національному університеті імені Петра Могили за адресою: 54003, м. Миколаїв, вул. 68-ми Десантників, 10.

З магістерською кваліфікаційною роботою можна ознайомитися в бібліотеці Чорноморського національного університету імені Петра Могили за адресою: 54003, м. Миколаїв, вул. 68-ми Десантників, 10.

Автореферат представлений «16» лютого 2021 р.

Секретар
екзаменаційної комісії,
к.пед.н., доцент Н. М. Болюбаш

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність дослідження визначається зростаючою тенденцією до пришвидшення процесів розробки та випуску програмних продуктів, що призводить до наявності у них великої кількості програмних дефектів.

Метою магістерської кваліфікаційної роботи є виконання дослідження з метою виявлення найбільш проблематичних секцій досліджуваних IT-проектів. Отримані дані можуть бути використані для подальшої оптимізації роботи команд розробників.

Об'єктоми дослідження виступатимуть програмні продукти п'яти реальних IT-проектів.

Предметом дослідження виступатимуть дефекти об'єктів дослідження.

Практичне значення даної магістерської кваліфікаційної роботи полягає у отриманні вичерпної та систематизованої інформації, на основі якої відбудеться перерозподіл поточних пріоритетів команд розробників з метою вирішення та запобігання найбільш серйозних дефектів програмних продуктів. Як наслідок - число потенційних помилок програмних застосунків буде скорочено, а разом із ними - час та кошти, що витрачаються на їх вирішення.

Доцільність застосування методу ABC-аналізу для оптимізації роботи команд розробників було розглянуто та опубліковано у тезах Всеукраїнської науково-практичної конференції молодих вчених, аспірантів і студентів «Інтелектуальні інформаційні системи – 2021» у секції «Системний аналіз, моделі і засоби підтримки прийняття рішень».

Магістерська кваліфікаційна робота складається із вступу, 5 розділів, висновків, додатку. Загальний обсяг роботи складає 99 сторінок, 29 рисунків, 15 таблиць та 46 посилань на літературні джерела.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі магістерської кваліфікаційної роботи обґрунтовано актуальність обраної теми, сформульовано мету і задачі дослідження, визначено предмет та об’єкт дослідження.

У першому розділі наведено огляд теоретичних засад збору та аналізу даних помилок продукції IT-проектів. Аналіз наявних аналогів та публікацій показав, що вдосконалені існуючі та нові рішення у даному напрямку дозволять оптимізувати роботу команд розробників, що у свою чергу вплине на якість кінцевих програмних продуктів. З урахуванням проведеного аналізу сформовано постановку задачі.

У другому розділі було розроблено спеціальну модульну систему для автоматизації збору статистичних даних помилок п’яти IT-проектів.

Збір статистичних даних відбувався за наступним сценарієм:

1. Визначившись з об’єктом дослідження, користувач запускає на виконання відповідний модуль.
2. Веб-драйвер переходить на головну сторінку багтрекера та проходить авторизацію у системі.
3. Веб-драйвер переходить на сторінку розширеного пошуку.
4. Відповідно до об’єкту дослідження, веб-драйвер налаштовує фільтри пошуку.
5. Один за одним, веб-драйвер переглядає сторінки програмних недоліків, збирає необхідну інформацію та записує її до БД.

Рис. 2.1. Блок-схема роботи модуля

Збір статистичних даних окремо взятої програмної помилки відбувався за наступним сценарієм:

1. Вибір дефекту із списку наявних дефектів.
2. Збір ключа та короткого опису програмного недоліку.
3. Перехід на сторінку дефекта у Jira.
4. Збір часу, ступеня критичності та поточного статусу програмної помилки.
5. Запис зібраних даних до БД.

Рис. 2.2. Блок-схема збору даних дефекта

Оскільки, на головній сторінці програмної помилки час відсутній, то життєвий цикл дефекту розраховувався за наступною формулою:

$$L_l = \sum S_l \tag{2.1}$$

де L_l – ¿ довжина життєвого циклу програмної помилки;

S_l – ¿тривалість виконання саб-таску.

Збір даних про життєвий цикл програмного недоліку відбувався за наступним сценарієм:

1. Перехід на сторінку дефекта у Jira.
2. Якщо саб-таски відсутні, то у якості тривалості життєвого циклу програмної помилки записується стандартне значення тривалості саб-таску – 30 хвилин.
3. Якщо саб-таски присутні, то веб-драйвер переходить на сторінку саб-таску.

4. Якщо час виконання саб-таску відсутній, то у якості даної величини записується стандартне значення тривалості саб-таску.
5. Після опрацювання останнього наявного саб-таску, тривалість життєвого циклу дефекта розраховується за формулою, що наведено вище.

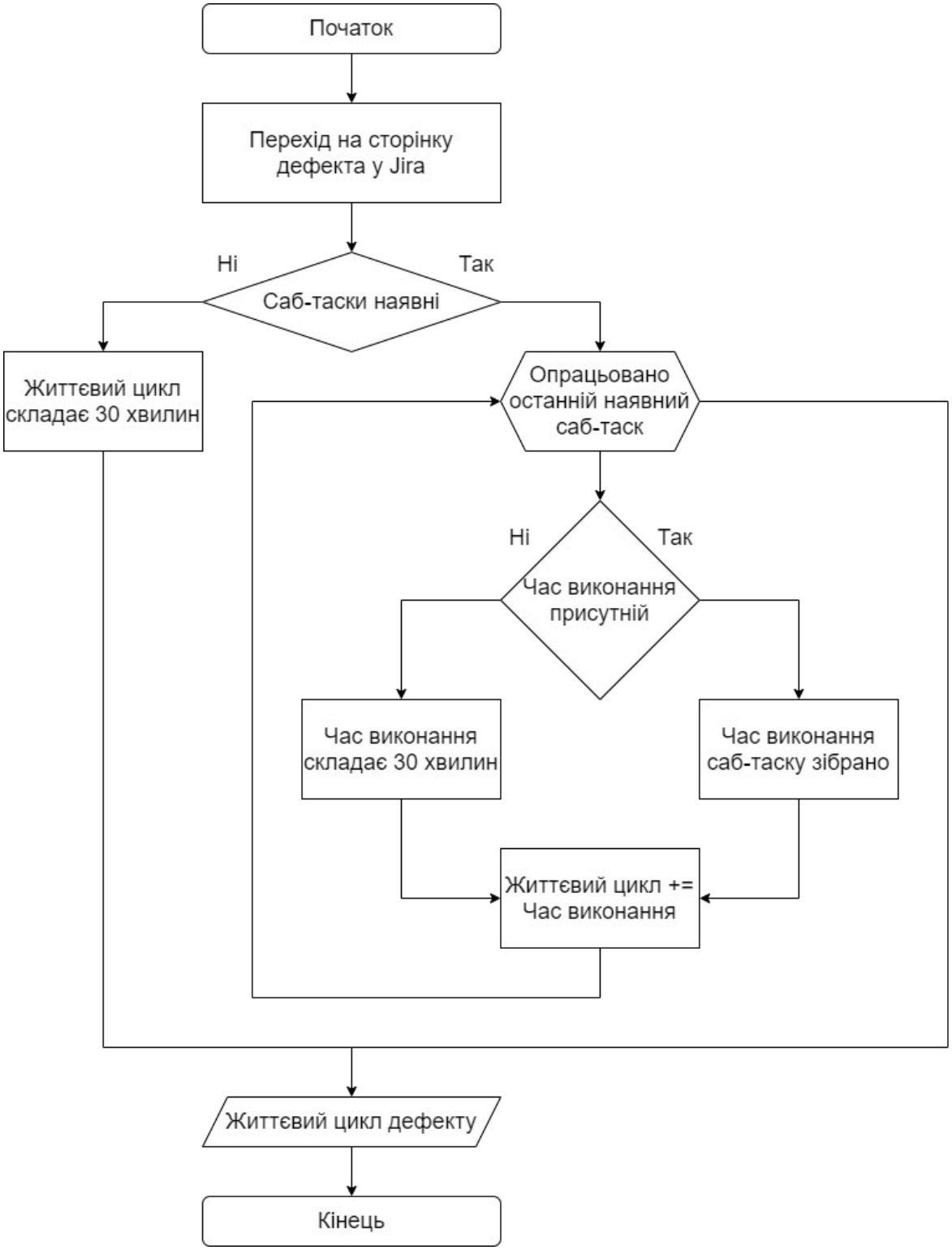

Рис. 2.3. Блок-схема збору даних про життєвий цикл дефекта

Позаяк, життєвий цикл програмної помилки буде безпосередньо використовуватися для ABC-аналізу, було прийняте рішення про конвертацію одиниць часу до числового формату для зручності подальших розрахунків. 1 година часу буде прирівняна до числа 1. Конвертація хвилин, днів та тижнів буде відбуватися згідно з наступними формулами:

$$S_l = \frac{m}{60} \tag{2.2}$$

$$S_l = d \cdot 8 \tag{2.3}$$

$$S_l = w \cdot 5 \cdot 8 \tag{2.4}$$

де $S_l -$ ¿ час, що витрачено на саб-таск,

$m -$ ¿кількість хвилин,

$d -$ ¿кількість днів,

$w -$ ¿кількість тижнів.

Так як тривалість робочого дня – 8 годин, а робочого тижня – 5 днів, то кількість днів та тижнів перемножувались на відповідні величини.

Визначившись з бізнес-логікою модульної системи для збору статистичних даних, було спроектовано базу даних (БД) для зберігання зібраної інформації. Для організації реляційних таблиць БД було обрано схему «зірка».

Кожна таблиця БД відповідає окремо взятому проекту та містить наступні поля:

1. Ключ (унікальний ідентифікатор) дефекту у системі.
2. Короткий опис програмного недоліку.
3. Тривалість життєвого циклу дефекту.
4. Ступінь критичності програмної помилки.
5. Поточний статус дефекту у системі.

Центром «зірки» виступає таблиця, що містить лише одне поле – ступінь критичності програмного недоліку. Дане рішення викликане тим, що це єдине поле, яке наявне в усіх таблицях «променів» і містить однакові значення в незалежності від таблиці.

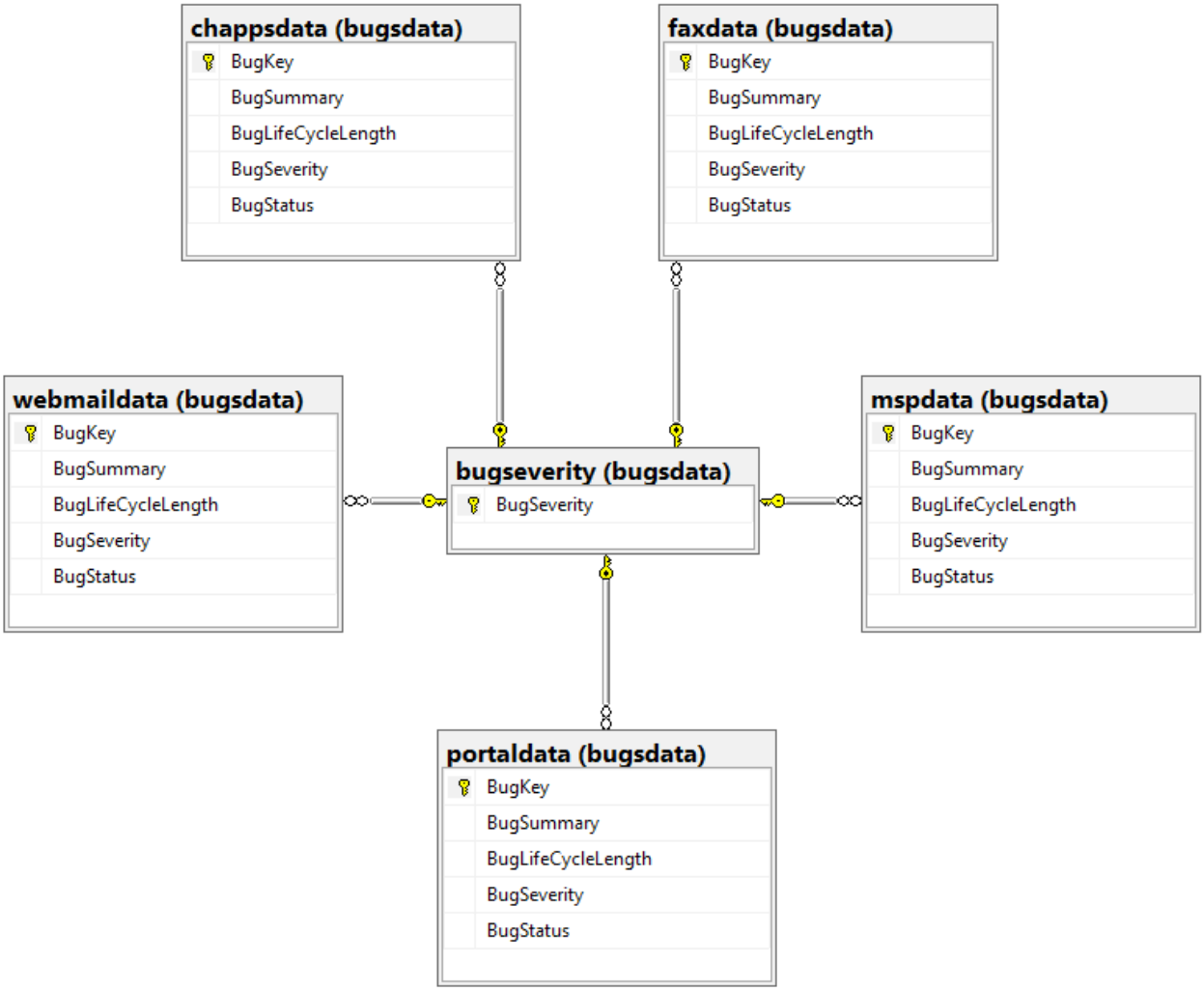

Рис. 2.7. ER діаграма створеної БД

Результатом застосування розробленої системи є зібрані статистичні дані, які сформували репрезентативні вибірки для подальшого дослідження та були записані до БД.

В третьому розділі було розроблено програмний застосунок (ПЗ) для виконання ABC-аналізу статистичних даних з урахуванням особливостей предмету дослідження.

Результати досліджень представлені у вигляді таблиць. Ці таблиці містять дані як у числовому, так і у відсотковому вигляді. У відсотках представлене відношення часу груп до загальної кількості витраченого часу.

Таблиця 3.1.

Результати аналізу статистичних даних дефектів IT-проектів

Назва Проекту	Група A		Група B		Група C	
Cloud Hosting Applications	11	16,4%	52	20,2%	713	63,4%
Online Fax	5	33,6%	11	33,2%	107	33,2%
Modular Sales Platform	17	46,1%	6	7,3%	309	46,6%
Cloud Portal	8	26,8%	21	30,2%	192	43%
Webmail	14	43,3%	23	26%	302	30,7%

Результати дослідження показали, що на виправлення дефектів групи A команди розробників витрачали досить велику кількість часу. Так, програмні помилки групи A проектів Online Fax, Modular Sales Platform та Webmail зайняли 33,6%, 46,1% та 43,3% часу від загального витраченого часу на виправлення усіх програмних недоліків даних проектів.

Таблиця 3.2.

Найбільш проблематичні секції IT-проектів

Назва проекту	Секції проекту
Cloud Hosting Applications	1. Менеджер доменів: – авто-оновлення домену; – вкладки менеджеру доменів; – нейм-сервери; – відображення доменів з певними доменами верхнього рівня. 2. Менеджер електронних адрес: – запуск менеджера електронних адрес; – авторизація у Webmail. 3. Список електронних адрес: – групове виділення електронних адрес; – відображення списку підписників у певних браузерах.
Online Fax	1. Оновлений до PHP 7+ факс-портал. 2. Налаштування факс-порталу: – загальні; – безпека. 3. Робота з повідомленнями:

Назва проекту	Секції проекту
	– завантаження зображень при написані повідомлення; – відображення надісланих повідомлень.
Modular Sales Platform	1. Платформа для продажів першої компанії: – вибір тривалості сплаченої послуги (місяць/рік); – ціна SSL сертифікату; – авторизація у порталі; – відображення даних секцій для зворотного зв'язку; – банери головної сторінки; – відображення цін на смартфонах; – кошик. 2. Платформа для продажів другої компанії: – кошик; – користувацький інтерфейс; – секція продуктів онлайн факсу. 3. Платформа для продажів третьої компанії: – створення аккаунтів та авторизація; – скидання паролю. 4. Інші компанії: – реєстрація домену; – завершення покупки.
Cloud Portal	1. Авторизація: – вибір мови перед авторизацією; – скидання паролю. 2. Вибір доменів. 3. Права доступу користувачів. 4. Управління придбаними продуктами.
Webmail	1. Авторизація 2. Контакти: – імпорт контактів; – використання робочої пошти контактів. 3. Перегляд отриманих повідомлень: – завантаження отриманих повідомлень; – відображення вкладених об'єктів. 4. Написання повідомлень. – відображення вкладених об'єктів; – валідація введених даних при використанні певних шаблонів; 5. Прив'язка до менеджеру електронних адрес. 6. Застосування користувацьких конфігурацій.

Таблиця 3.3.

Результати аналізу статистичних даних дефектів IT-проектів з використанням користувацьких коефіцієнтів критичності

Назва Проекту	Група A		Група B		Група C	
Cloud Hosting Applications	10	16%	64	22,8%	702	61,2%
Online Fax	5	33,6%	11	33,5%	107	32,9%
Modular Sales Platform	13	42,4%	11	14,1%	308	43,5%
Cloud Portal	8	27,5%	21	29,6%	192	42,9%
Webmail	15	46,2%	22	24,8%	302	29%

Результати виконаного дослідження показали незначні відхилення від попередніх даних у деяких IT-проектах. Так, група A проекту Cloud Hosting Applications зменшилась з 11 до 10, Modular Sales Platform – з 17 до 13, а у Webmail – збільшилась з 14 до 15.

Таблиця 3.4.

Найбільш проблематичні секції IT-проектів

Назва проекту	Секції проекту
Cloud Hosting Applications	1. Менеджер доменів – авто-оновлення домену; – вкладки менеджеру доменів; – нейм-сервери; – відображення доменів з певними доменами верхнього рівня. 2. Менеджер електронних адрес – запуск менеджера електронних адрес; – авторизація у Webmail. 3. Списки електронних адрес: – відображення списку підписників у певних браузерах. 4. Менеджер FTP: – оновлення паролю для певних користувачів.
Online Fax	1. Оновлений до PHP 7+ факс-портал. 2. Налаштування факс-порталу: – загальні;

Назва проекту	**Секції проекту**
	– безпека. 3. Робота з повідомленнями: – завантаження зображень при написані повідомлення; – відображення надісланих повідомлень.
Modular Sales Platform	1. Платформа для продажів першої компанії: – вибір тривалості сплаченої послуги (місяць/рік); – ціна SSL сертифікату; – авторизація у порталі; – відображення даних секцій для зворотного зв’язку; – банери головної сторінки; – відображення цін на смартфонах; – кошик. 2. Платформа для продажів третьої компанії: – створення аккаунтів та авторизація; – скидання паролю. 3. Інші компанії: – реєстрація домену; – завершення покупки.
Cloud Portal	1. Авторизація: – вибір мови перед авторизацією; – скидання паролю. 2. Вибір доменів. 3. Права доступу користувачів. 4. Управління придбаними продуктами.
Webmail	1. Авторизація 2. Контакти: – імпорт контактів; – використання робочої пошти контактів. 3. Перегляд отриманих повідомлень: – завантаження отриманих повідомлень; – відображення вкладених об’єктів. 4. Написання повідомлень. – відображення вкладених об’єктів; – валідація введених даних при використанні певних шаблонів; 5. Прив’язка до менеджеру електронних адрес. 6. Застосування користувацьких конфігурацій. 7. Кеш зібраних електронних адрес.

Таблиця 3.5.

Результати аналізу статистичних даних дефектів IT-проектів з урахуванням лише виправлених програмних помилок

Назва Проекту	Група A		Група B		Група C	
Cloud Hosting Applications	7	16,9%	24	19,6%	369	63,5%
Online Fax	1	12,3%	8	43,2%	50	44,6%
Modular Sales Platform	15	43,6%	7	7,5%	295	48,9%
Cloud Portal	6	23,1%	16	29,6%	168	47,3%
Webmail	2	15,9%	11	37,3%	79	46,8%

Результати виконаного дослідження показали значні відхилення від усіх попередньо отриманих даних. Так, група A проекту Cloud Hosting Applications зменшилась до 7, Online Fax – до 1, Modular Sales Platform – збільшилась з 13 до 15, однак це все ще нижче за перші отримані дані – 17, Cloud Portal – зменшилась до 6, а Webmail – до 2. Ще одне цікаве спостереження – кількість невиправлених недоліків перевищує у 2 та більше разів на проектах Cloud Hosting Applications, Online Fax та Webmail. Ймовірно, такі результати пов’язані з тим, що ці проекти мають складну архітектуру, через що їх досить важко підтримувати у належному стані. Втім, не варто виключати й людського фактору – команди розробників можуть не встигати робити необхідні виправлення через брак ресурсів.

Таблиця 3.6.

Найбільш проблематичні секції IT-проектів

Назва проекту	Секції проекту
Cloud Hosting Applications	1. Менеджер доменів – авто-оновлення домену; – вкладки менеджеру доменів; – нейм-сервери; – відображення доменів з певними доменами верхнього рівня. 2. Менеджер електронних адрес – запуск менеджера електронних адрес; – авторизація у Webmail.

Назва проекту	Секції проекту
	3. Списки електронних адрес: – групове виділення електронних адрес.
Online Fax	1. Оновлений до PHP 7+ факс-портал.
Modular Sales Platform	1. Платформа для продажів першої компанії: – вибір тривалості сплаченої послуги (місяць/рік); – ціна SSL сертифікату; – авторизація у порталі; – відображення даних секцій для зворотного зв'язку; – банери головної сторінки; – відображення цін на смартфонах; – кошик. 2. Платформа для продажів другої компанії: – кошик; – користувацький інтерфейс; – секція продуктів онлайн факсу. 3. Платформа для продажів третьої компанії: – створення аккаунтів та авторизація; – скидання паролю. 4. Інші компанії: – реєстрація домену; – завершення покупки.
Cloud Portal	1. Авторизація: – вибір мови перед авторизацією; – скидання паролю. 2. Права доступу користувачів. 3. Управління придбаними продуктами.
Webmail	1. Контакти: – використання робочої пошти контактів. 2. Перегляд отриманих повідомлень: – завантаження отриманих повідомлень.

Таблиця 3.7.

Результати аналізу статистичних даних дефектів IT-проектів з урахуванням усіх наявних фільтрів

Назва Проекту	Група A		Група B		Група C	
Cloud Hosting Applications	6	15,8%	27	20,5%	367	63,7%

Online Fax	1	11,2%	9	47,9%	49	40,9%
Modular Sales Platform	10	37,8%	11	15,6%	296	46,6%
Cloud Portal	7	26,7%	18	29,1%	165	44,2%
Webmail	2	18,2%	12	37,8%	78	44%

Результати виконаного дослідження показали незначні відхилення від даних, що були отримані в результаті відсіювання усіх невиправлених програмних помилок. Так, група А проекту Cloud Hosting Applications зменшилась з 7 до 6, Modular Sales Platform – з 15 до 10, Cloud Portal – збільшилась з 6 до 7.

Таблиця 3.8.

Найбільш проблематичні секції IT-проектів

Назва проекту	**Секції проекту**
Cloud Hosting Applications	1. Менеджер доменів: – авто-оновлення домену; – вкладки менеджеру доменів; – нейм-сервери; – відображення доменів з певними доменами верхнього рівня. 2. Менеджер електронних адрес: – запуск менеджера електронних адрес; – авторизація у Webmail.
Online Fax	1. Оновлений до PHP 7+ факс-портал.
Modular Sales Platform	1. Платформа для продажів першої компанії: – вибір тривалості сплаченої послуги (місяць/рік); – ціна SSL сертифікату; – авторизація у порталі; – відображення даних секцій для зворотного зв’язку; – банери головної сторінки; – відображення цін на смартфонах; – кошик. 2. Інші компанії: – реєстрація домену; – завершення покупки.

Назва проекту	**Секції проекту**
Cloud Portal	1. Авторизація: – вибір мови перед авторизацією; – скидання паролю. 2. Права доступу користувачів. 3. Управління придбаними продуктами.
Webmail	1. Контакти: – використання робочої пошти контактів. 2. Перегляд отриманих повідомлень: – завантаження отриманих повідомлень.

У методичній частині магістерської кваліфікаційної роботи було розроблено ряд практичних робіт, метою яких є ознайомлення студентів зі сферою тестування програмного забезпечення, формування в них загального уявлення та базових навичок для виконання функціонального тестування.

У спеціальній частині магістерської кваліфікаційної роботи з «Охорони праці та безпеки життєдіяльності» було встановлено, що приміщення, у якому працює виконавець, не задовольняє умовам діючого законодавства в сфері охорони праці. Так, було встановлено, що в приміщенні не вистачає світла, а також достатнього рівня повітропродуктивності системи вентиляції, що є необхідною умовою для компенсації вологонадлишків у виробничому приміщенні.

ЗАГАЛЬНІ ВИСНОВКИ

У ході виконання магістерської роботи, було виконано усі поставлені задачі, а саме:

1. Розкрито теоретичні засади збору та аналізу даних програмних помилок. У ході виконання даного завдання, було виявлено та доведено, що дефекти програмних продуктів відіграють значну роль у життєвому циклі будь-якого програмного продукту. Про значимість даної ролі говорить незліченна кількість наукових публікацій та програмних засобів, що спрямовані на зменшення кількості програмних помилок та частоту їх виникнення. Вдосконалені існуючі та нові рішення у даному напрямку

дозволять оптимізувати роботу команд розробників, що у свою чергу вплине на якість кінцевих програмних продуктів.

2. Виконано збір статистичних даних та сформовано репрезентативні вибірки. У ході виконання даного завдання, було розроблено спеціальну модульну систему для автоматизації збору статистичних даних помилок п’яти IT-проектів. Результатом застосування даної системи є зібрані статистичні дані, які сформували репрезентативні вибірки для подальшого дослідження.
3. Розроблено програмне забезпечення для виконання ABC-аналізу з урахуванням особливостей предмету дослідження та проаналізовано статистичні дані кожного об’єкту дослідження з формуванням відповідних висновків. У ході виконання даних завдань, було розроблено цільове програмне забезпечення та виконано аналіз статистичних даних помилок п’яти IT-проектів. У результаті проведеного аналізу, було виявлено їх найбільш проблематичні сектори.

Окрім усього вище переліченого, було виконано методичну та спеціальну частини магістерської роботи.

У ході виконання методичного розділу магістерської роботи, було розроблено ряд практичних робіт, метою яких є ознайомлення студентів зі сферою тестування програмного забезпечення, формування в них загального уявлення та базових навичок для виконання функціонального тестування.

У ході виконання спеціальної частини магістерської роботи було встановлено, що приміщення, у якому працює виконавець, не задовольняє умовам діючого законодавства в сфері охорони праці. Так, було встановлено, що в приміщенні не вистачає світла, а також достатнього рівня повітропродуктивності системи вентиляції, що є необхідною умовою для компенсації вологонадлишків у виробничому приміщенні.

У підсумку, можна стверджувати, що мета даної магістерської роботи була досягнена у повному обсязі.

АНОТАЦІЯ

Казакевич Дмитро Андрійович. Статистичний аналіз вибірки найбільш поширених помилок в IT-проектах для оптимізації роботи команд розробників. – На правах рукопису.

Магістерська кваліфікаційна робота на здобуття освітньої кваліфікації «Магістр комп'ютерних наук». – Чорноморський національний університет імені Петра Могили, Миколаїв, 2021.

Дана магістерська кваліфікаційна робота присвячена статистичному аналізу вибірок програмних помилок продукції IT-проектів для оптимізації роботи команд розробників.

Метою виконання статистичного аналізу дефектів є оптимізація роботи команд розробників.

Об'єктами дослідження виступають програмні продукти п'яти IT-проектів.

Предметом дослідження виступають дефекти об'єктів дослідження.

Фахова частина магістерської кваліфікаційної роботи складається з наступних розділів: дослідження предметної області; збір статистичних даних та формування репрезентативних вибірок; розробка програмного забезпечення для аналізу статистичних даних.

Задачі, які були виконані в процесі роботи:

- розкрито теоретичні засади збору та аналізу даних програмних помилок;
- виконано збір статистичних даних та сформовано репрезентативні вибірки;
- розроблено програмне забезпечення для виконання ABC-аналізу з урахуванням особливостей предмету дослідження;
- виконано аналіз статистичних даних кожного об'єкту дослідження та сформульовано відповідні висновки.

В спеціальній частині магістерської кваліфікаційної роботи з «Охорони праці та безпеки у надзвичайних ситуаціях» проведено аналіз освітлення та вентиляційного режиму в виробничому приміщенні.

У методичній частині розроблено практичні роботи з курсу: «Основи тестування програмного забезпечення. Функціональне тестування».

Робота складається з 99 сторінок, 29 рисунків, 15 таблиць та 46 посилань на літературні джерела.

Ключові слова: *статистика, ABC-аналіз, програмний дефект, оптимізація.*

ABSTRACT

Kazakevych Dmytro. Statistical analysis of the most common errors of IT-projects for work optimization of development teams – On the rights of the manuscript.

Master's qualification work for obtaining an educational qualification "Master of Computer Science". – Petro Mohyla Black Sea National University, Mykolaiv, 2020.

The master's qualification work is devoted to statistical analysis of the most common errors of IT-projects for work optimization of development teams.

The purpose of performing statistical analysis of defects is to optimize the work of development teams.

The objects of the study are the software products of five IT projects.

The subject of research are defects of objects of research.

The professional part of the master's qualification work consists of the following sections: research of the subject area; collection of statistical data and formation of representative samples; development of software for statistical data analysis.

Tasks that were performed in the process:

- theoretical principles of data error data collection and analysis are revealed;
- statistical data were collected and representative samples were formed;
- developed software for ABC analysis, taking into account the characteristics of the subject of study;
- the analysis of statistical data of each object of research is executed and the corresponding conclusions are formulated.

In the special part of the master's qualification work on "Occupational Health and Safety in Emergencies" analysis of lighting and ventilation in the production room was performed.

In the methodical part practical works from a course: «Fundamentals of software testing. Functional testing» have been developed.

The work consists of 99 pages, 29 figures, 15 tables and 46 references.

Keywords: *statistics, ABC analysis, software defect, optimization.*