

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ
ФАКУЛЬТЕТ КОМП'ЮТЕРНО-ІНФОРМАЦІЙНИХ СИСТЕМ І ПРОГРАМНОЇ
ІНЖЕНЕРІЇ
КАФЕДРА КОМП'ЮТЕРНИХ НАУК

ПАНАСЮК ІГОР ВОЛОДИМИРОВИЧ

УДК 004.9

**ОРГАНІЗАЦІЯ ХМАРНОЇ ПЛАТФОРМИ «РОЗУМНОГО МІСТА» ДЛЯ
УНИКНЕННЯ ПОВТОРІВ КОЛЕКЦІЙ ДАНИХ**

126 – Інформаційні системи та технології

Автореферат

дипломної роботи на здобуття освітнього ступеня «магістр»

Тернопіль
2019

Роботу виконано на кафедрі комп'ютерних наук Тернопільського національного технічного університету імені Івана Пулюя Міністерства освіти і науки України

Керівник роботи: доктор технічних наук,
професор кафедри комп'ютерних наук
Пасічник Володимир Володимирович,
Тернопільський національний технічний університет
імені Івана Пулюя,

Рецензент: кандидат технічних наук,
доцент кафедри програмної інженерії
Михалик Дмитро Михайлович,
Тернопільський національний технічний університет
імені Івана Пулюя,

Захист відбудеться 24 грудня 2019 р. о 9⁰⁰ годині на засіданні
екзаменаційної комісії № 29 у Тернопільському національному технічному
університеті імені Івана Пулюя за адресою: 46001, м. Тернопіль, вул. Руська, 56,
навчальний корпус №1, ауд. 702

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми роботи. Провівши аналіз поточного стану досліджень було виявлено ряд аспектів які потрібно додатково вивчити. Наприклад, як ефективніше використовувати внутрішні характеристики даних, пропонувати більш ефективні технології уникнення повторів, як підвищити надійність систем уникнення повторів і як поліпшити дедуплікацію даних. Продуктивність технології, інтеграція різних існуючих технологій в проектах «розумних міст», підвищення гнучкості, масштабованості і адаптованості технології уникнення повторів є актуальним напрямком сучасних досліджень.

Мета і задачі дослідження. Підвищення якості сервісів інформаційно-технологічної платформи «розумного міста» та зменшення витрат на оренду її хмарних обчислювальних ресурсів.

Об'єкт, методи та джерела дослідження: Підсистема для хмарного зберігання та опрацювання даних в проектах класу «розумне місто».

Методи аналітичного опрацювання наукових публікацій та джерел. Методи системного аналізу. Методи видобування даних. Методи системного та логічного програмування. Методи дедублікації.

Як інформаційні джерела використовуються наукові публікації та інтернет джерела.

Наукова новизна отриманих результатів:

Описано процес запуску «spark RDD» в середовищі «Spark» для проектів класу «розумне місто».

Практичне значення отриманих результатів.

Реалізовано метод уникнення повторів для потреб проектів класу «розумне місто».

Апробація результатів магістерської роботи проведена на двох наукових конференціях з публікацією тез доповідей.

Структура роботи. Робота складається з розрахунково-пояснювальної записки та графічної частини. Розрахунково-пояснювальна записка складається з вступу, 7 частин, висновків, переліку посилань та додатків. Обсяг роботи: розрахунково-пояснювальна записка – 115 арк. формату А4, графічна частина – 12 слайдів презентації.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** проведено огляд сучасного стану досліджень. Сформовано тему, мету та завдання роботи.

В **першому розділі** дипломної роботи досліджено інноваційні проекти класу «розумне місто». Виконано огляд та проаналізовано способи зменшення вартості зберігання даних в інноваційних проектах «розумних міст». Розкрито зміст інформаційної технології «Дедуплікація» даних в хмарних проектах «розумних міст».

В **другому розділі** дипломної роботи досліджено порівняння способів «уникнення повторів» застосованих в проектах «розумних міст». Запропоновано

архітектуру проектованої хмарної платформи «розумного міста» для уникнення повторів колекцій даних. Проаналізовано ключові особливості файлової системи «OpenDedup» з автоматичним об'єднанням дублікатів даних. Досліджено архітектуру екосистеми «Hadoop» в контексті її використання для проектів «розумних міст». Подано Огляд «HBase» для проектів «розумних міст» та описано реалізацію методу уникнення повторів даних в проектах класу «розумне місто». В розділі розроблений алгоритм MapReduce для уникнення повторів великих даних. Описано підготовку та налаштування операційного середовища «Hadoop».

В третьому розділі дипломної роботи описано процес запуску кластера «Hadoop» в «Docker»-контейнерах для потреб «розумних міст». Розглянуто процес запуску «spark RDD» в середовищі «Spark» для проектів класу «розумне місто». Висвітлено практичну реалізацію методу уникнення повторів для потреб проектів класу «розумне місто».

В розділі «**Спеціальна частина**» описано процеси виконання кластерних експериментів.

В розділі «**Обґрунтування економічної ефективності**» розраховано основні техніко-економічні показники проведених досліджень.

В розділі «**Охорона праці та безпека в надзвичайних ситуаціях**» описано основні особливості стандарту OHSAS 18001 щодо ведення та управління документацією з охорони праці та розглянуто заходи щодо відвернення пожежі на робочих місцях користувачів ПК. Висвітлено питання оцінки події, що сталася або може статися у прогнозований термін, та визначення ступеня реагування на відповідному рівні управління. Описано процес функціонування державної системи спостереження, збирання, оброблення та аналізу інформації про стан довкілля під час надзвичайних ситуацій мирного та воєнного часу.

В розділі «**Екологія**» описано методи узагальнення екологічної інформації та отримання енергії за рахунок альтернативних джерел.

У загальних висновках щодо дипломної роботи описано прийняті в дипломній роботі освітнього рівня «Магістр» наукові та технічні рішення і організаційно-технічні заходи, які забезпечують виконання завдання на проектування; оригінальні технічні рішення, прийняті автором в процесі роботи; технічні рішення роботи, які можуть бути впроваджені у виробництво;

В додатках до пояснювальної записки приведено ксерокопії тез доповідей.

В графічній частині подано тему, мету, об'єкт та предмет дослідження. Подано завдання до дипломної роботи. Описано «розумне місто», структура «Opendedup» в проектах класу «Розумне місто», узагальнена структура екосистеми «Hadoop», взаємодія складових сутностей «HDFS», запис даних «розумного міста», алгоритм уникнення повторів, тестування платформи. Описані висновки.

ВИСНОВКИ

В процесі виконання дипломної роботи освітнього рівня «магістр» була виконана організація хмарної платформи «розумного міста» для уникнення повторів колекцій даних. В процесі аналізу предметної області:

–досліджено інноваційні проекти класу «розумне місто»;

–виконано огляд та проаналізовано способи зменшення вартості зберігання даних в інноваційних проектах «розумних міст»;

–розкрито зміст інформаційної технології «Дедуплікація» даних в хмарних проектах «розумних міст».

В другому розділі дипломної роботи:

–подано порівняння способів «уникнення повторів» застосованих в проектах «розумних міст»;

–запропоновано архітектуру проектованої хмарної платформи «розумного міста» для уникнення повторів колекцій даних;

–проаналізовано ключові особливості файлової системи «OpenDedup» з автоматичним об'єднанням дублікатів даних;

–досліджено архітектуру екосистеми «Hadoop» в контексті її використання для проектів «розумних міст»;

–подано Огляд «HBase» для проектів «розумних міст» та описано реалізацію методу уникнення повторів даних в проектах класу «розумне місто»;

–в розділі розроблений алгоритм MapReduce для уникнення повторів великих даних;

–описано підготовку та налаштування операційного середовища «Hadoop».

В третьому розділі дипломної роботи:

–описано процес запуску кластера «Hadoop» в «Docker»-контейнерах для потреб «розумних міст»;

–розглянуто процес запуску «spark RDD» в середовищі «Spark» для проектів класу «розумне місто»;

–висвітлено практичну реалізацію методу уникнення повторів для потреб проектів класу «розумне місто».

В розділі «Спеціальна частина» описано процеси виконання кластерних експериментів.

В розділі «Обґрунтування економічної ефективності» розраховано основні техніко-економічні показники проведених досліджень.

В розділі «Охорона праці та безпека в надзвичайних ситуаціях» описано основні особливості стандарту OHSAS 18001 щодо ведення та управління документацією з охорони праці та розглянуто заходи щодо відвернення пожежі на робочих місцях користувачів ПК. Висвітлено питання оцінки події, що сталася або може статися у прогнозований термін, та визначення ступеня реагування на відповідному рівні управління. Описано процес функціонування державної системи спостереження, збирання, оброблення та аналізу інформації про стан довкілля під час надзвичайних ситуацій мирного та воєнного часу.

В розділі «Екологія» описано методи узагальнення екологічної інформації та отримання енергії за рахунок альтернативних джерел.

СПИСОК ОПУБЛІКОВАНИХ АВТОРОМ ПРАЦЬ ЗА ТЕМОЮ РОБОТИ

1. Інтернет речей, «Великі дані» та аналітичне опрацювання в «розумному місті» / [Панасюк І.В. та ін.]. // Матеріали VII науково-технічної конфіції «Інформаційні моделі, системи та технології» Тернопільського національного технічного університету імені Івана Пулюя, (Тернопіль, 11 – 12 грудня 2019 р.). –

Тернопіль: Тернопільський національний технічний університет імені Івана Пулюя, 2019. – С. 45.

2. ПОЄДНАННЯ «ІНТЕРНЕТУ РЕЧЕЙ» ТА «ВЕЛИКИХ ДАНИХ» В ПРОЕКТАХ КЛАСУ «РОЗУМНЕ МІСТО» / [Панасюк І.В. та ін.]. // Збірник тез II міжнародної конференції молодих вчених та студентів, 4-5 грудня 2019 р. / За заг. ред. Н.В. Габрусєвої – Тернопіль: ТНТУ ім. І. Пулюя, 2019 – С. 35.

АНОТАЦІЯ

Дипломна робота присвячена організації хмарної платформи «розумного міста» побудованої на принципах уникнення повторів колекцій даних. В першому розділі дипломної роботи досліджено інноваційні проекти класу «розумне місто». Виконано огляд та проаналізовано способи зменшення вартості зберігання даних в інноваційних проектах «розумних міст». Розкрито зміст інформаційної технології «Дедуплікація» даних в хмарних проектах «розумних міст».

В другому розділі дипломної роботи розглянуто подано порівняння способів «уникнення повторів» застосованих в проектах «розумних міст». Запропоновано архітектуру проектованої хмарної платформи «розумного міста» для уникнення повторів колекцій даних. Проаналізовано ключові особливості файлової системи «OpenDedup» з автоматичним об'єднанням дублікатів даних. Досліджено архітектуру екосистеми «Nadoop» в контексті її використання для проектів «розумних міст». Подано огляд «HBase» для проектів «розумних міст» та описано реалізацію методу уникнення повторів.

Ключові слова: ДЕДУБЛІКАЦІЯ, ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ, ІНТЕРНЕТ, ІНТЕРФЕЙС, ПЛАТФОРМА, РОЗУМНЕ МІСТО, СЕРВЕР, ХМАРНІ ОБЧИСЛЕННЯ.

ANNOTATION

The diploma thesis is devoted to the research of methods and means of organization of fire safety systems for projects of the "smart city" class. In the process of domain analysis, the innovative concept of "smart city" is described. The information and technological concept of Big Data is described. The Big Data in the smart city and their use to implement the smart city service subsystems are considered. The second section of the thesis examines the processing of information in projects of the class "smart city". A description of the Surgard alarm message format used is provided. Messaging protocols are explored. Comparison of development tools is made. The list of characteristics of the fire safety subsystem has been formed. The subsystem architecture was designed. In the third section of the thesis the detailed architecture of the "smart city" fire safety subsystem is designed. Methods of testing the "smart city" fire safety subsystem have been developed. The sections "Special part", "Justification of economic efficiency", "Occupational health and safety", "Ecology" have been completed.

Keywords: DEDUBLICATION, INFORMATION TECHNOLOGIES, INTERNET, INTERFACE, PLATFORM, SMART PLACE, SERVER, CLOUD COMPUTING.