

# МЕТОД ТА ПРОГРАМНА РЕАЛІЗАЦІЯ БЛОЧНОГО ПОРІВНЯННЯ ФАЙЛІВ

Виконав: студент групи 1КІ-18м  
Гудименко Олександр Олександрович |  
Керівник: к. т. н., доц. каф. ОТ  
Савицька Людмила Анатоліївна

# Актуальність дослідження

- ▶ Порівняння файлів широко використовується в індустрії програмування, де ведеться спільна робота і необхідно часто відстежувати зміни, які вносять в робочі файли різні користувачі, тому швидкість при синхронізації каталогів є нагальною потребою - щоб протягом порівняння не були внесені зміни.
- ▶ Алгоритм роботи методу CRC 16 дозволяє звести порівняння до порівняння так званих хеш-кодів, які вираховуються як строкове представлення контрольних кодів для блоків, на які поділяється зміст файлу.

# Мета, об'єкт, предмет дослідження

- ▶ **Метою** дослідження магістерської кваліфікаційної роботи є пришвидшення порівняння достатньо великих обсягів даних
- ▶ **Об'єкт** дослідження - це сучасні процеси спільної роботи над проектами на підприємстві із використанням сучасних інформаційних технологій.
- ▶ **Предмет** дослідження - це методи та програмні засоби блочного аналізу та порівняння файлів.

# Наукова новизна одержаних результатів

- ▶ вперше запропоновано метод блочного порівняння файлів, без витрат на рядкове чи двійкове порівняння. Запропонований у даній роботі метод блочного порівняння файлів дозволяє пришвидшити достатньо великих обсягів даних;
- ▶ вдосконалено процес порівняння двох файлів для програми блочного порівняння файлів за рахунок застосування можливостей хеш-кодів із використанням методу CRC 16, який дозволяє відстежувати зміни у вмісті файлу;
- ▶ вдосконалено процеси підрахунку та порівняння хеш-кодів за методом CRC 16 без витрат на рядкове чи двійкове порівняння.

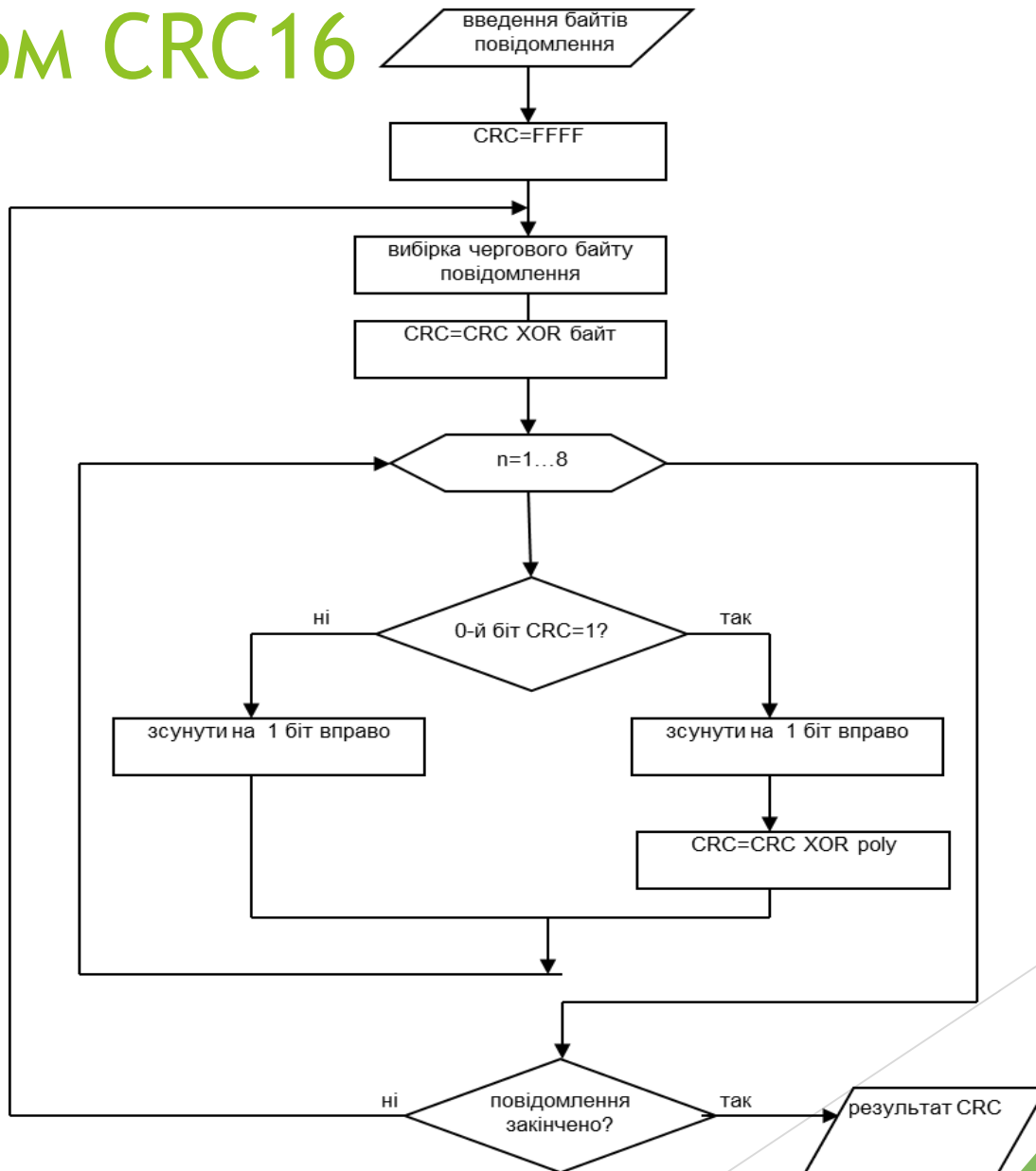
# Практичне значення одержаних результатів

- розроблено новий метод блочного порівняння файлів.
- вдосконалено процес порівняння двох файлів за рахунок застосування можливостей хеш-кодів із використанням методу CRC 16.
- розроблено алгоритм роботи процесів підрахунку та порівняння хеш-кодів за методом CRC 16, який дозволяє відстежувати зміни у вмісті файлу.

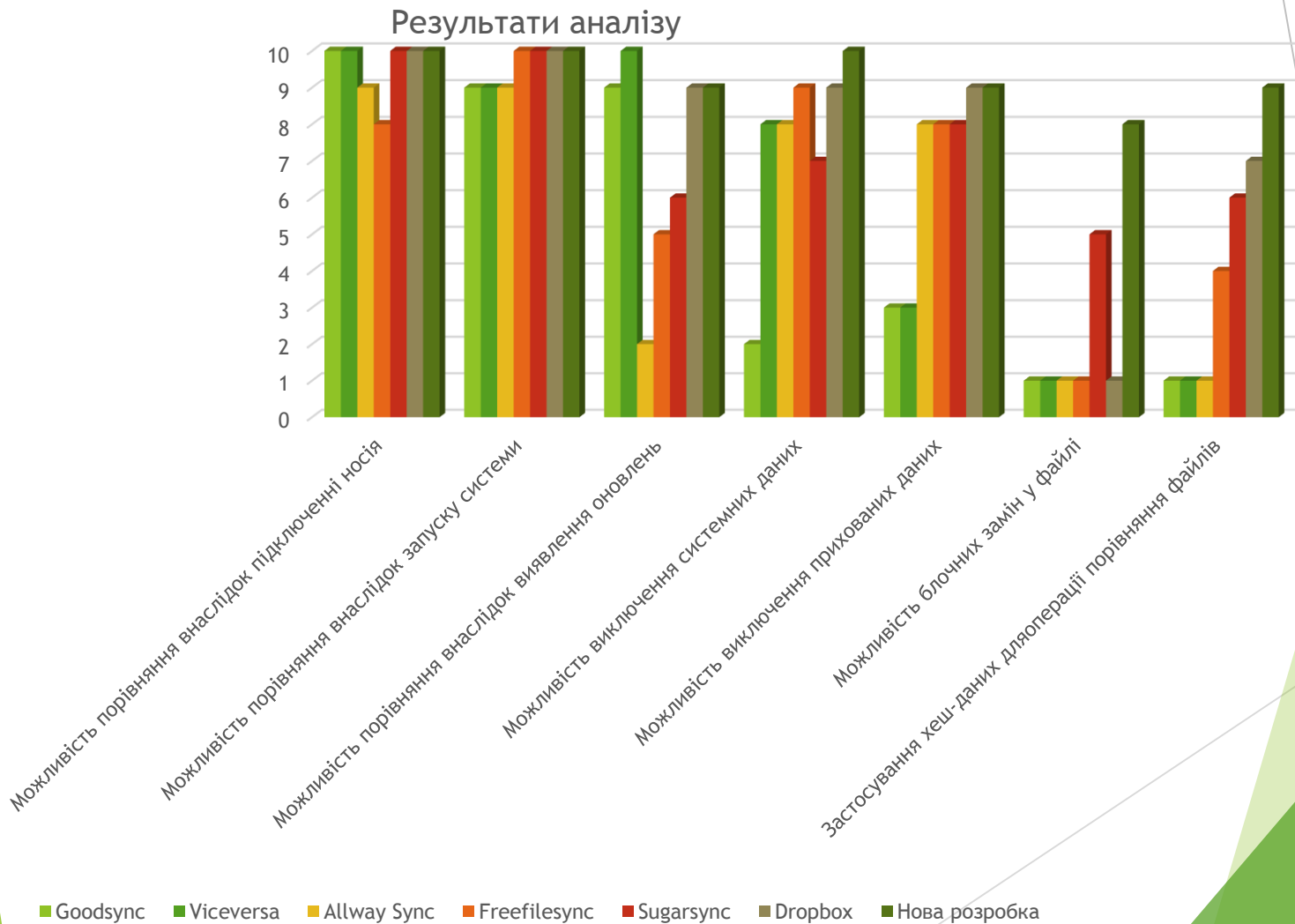
# Порівняння даних за CRC16

- ▶ метод CRC (Cyclic Redundancy Check) - один із методів розрахунку контрольної суми.
- ▶ в основі методу лежить представлення бітових послідовностей у вигляді многочленів з коефіцієнтами 0 або 1, де порядок біту відповідає степені доданка (починаючи з 0-го), а значення біту - його коефіцієнту.
- ▶ контрольна сума CRC використовується в багатьох мережних протоколах з різними утворюючими многочленами. Наприклад в Modbus використовується метод CRC16 з використанням 16-бітної контрольної суми. Популярним також є CRC32, який має 32-бітну контрольну суму.

# Алгоритм розрахунку КС за методом CRC16

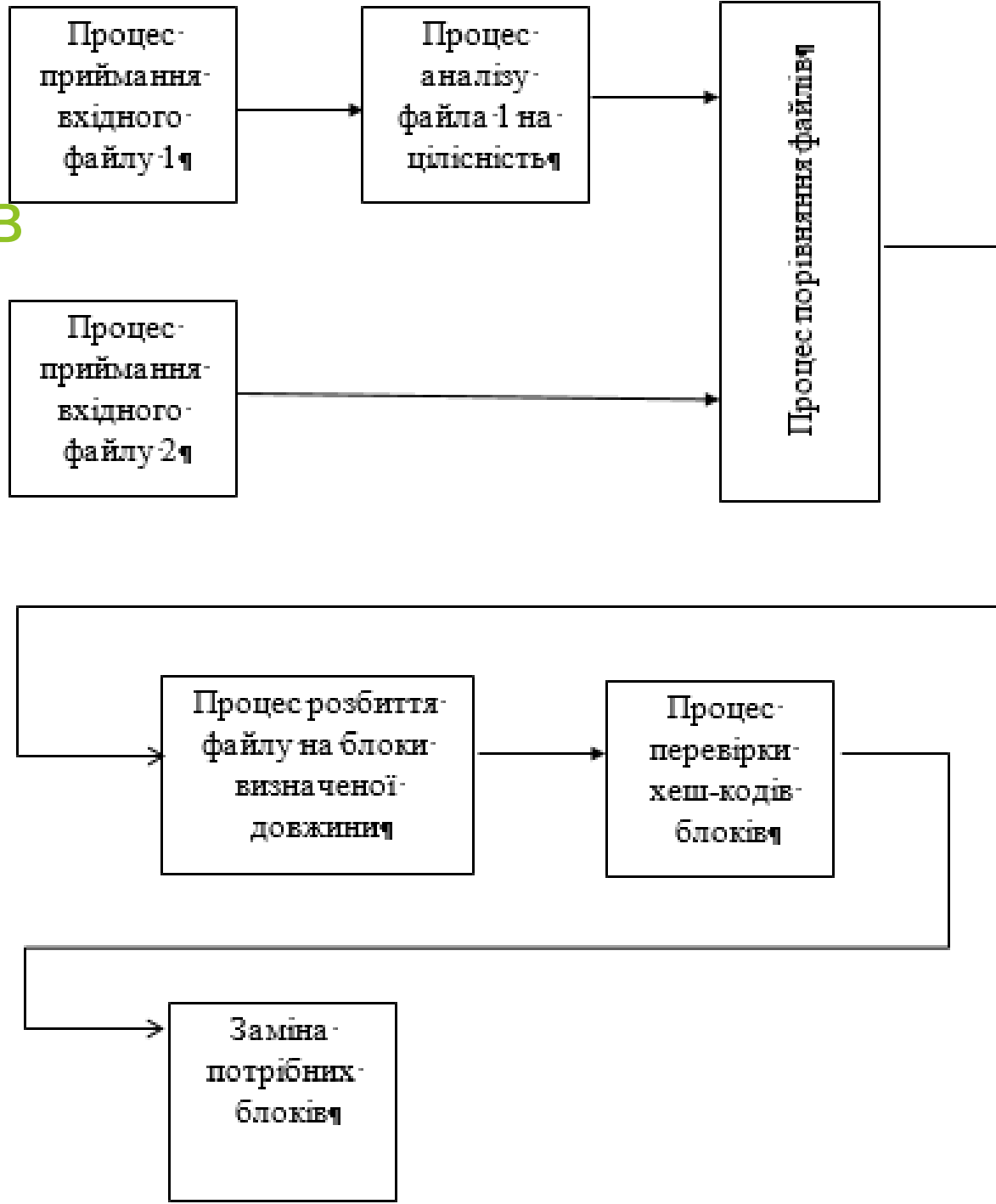


# Аналіз конкурентів

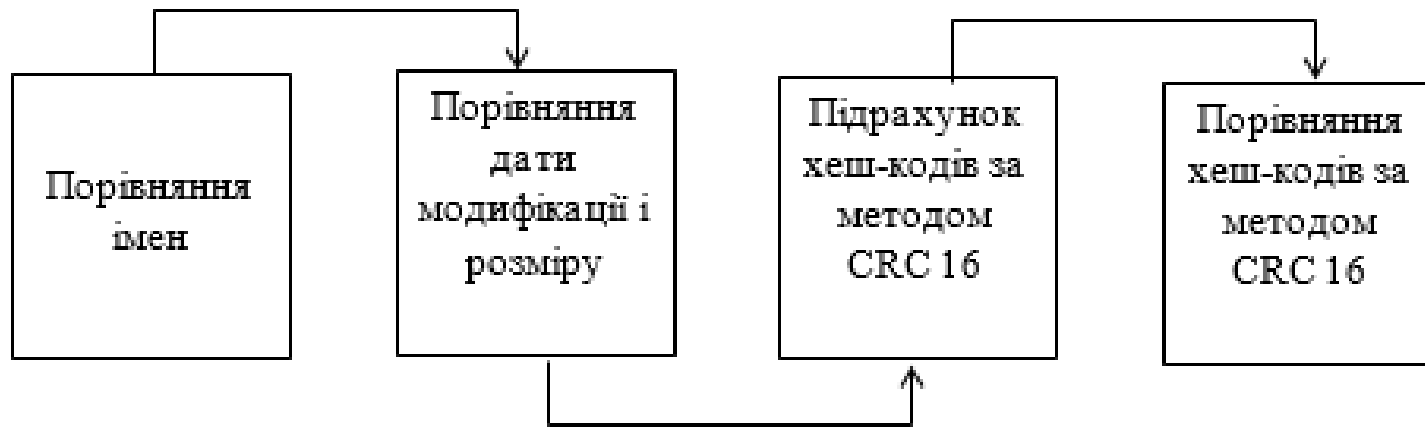




# Робота методу блочного порівняння файлів



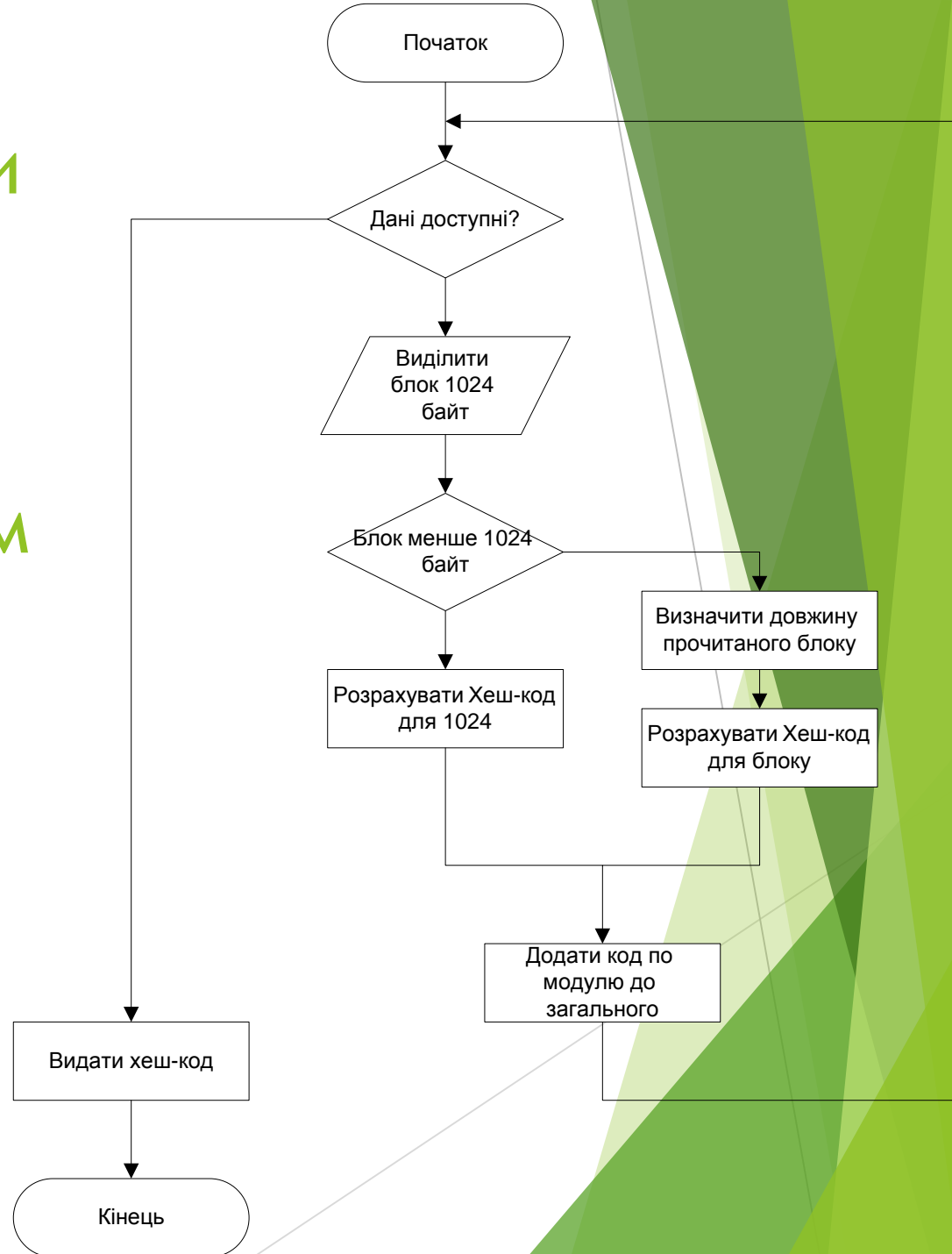
# Процес порівняння файлів



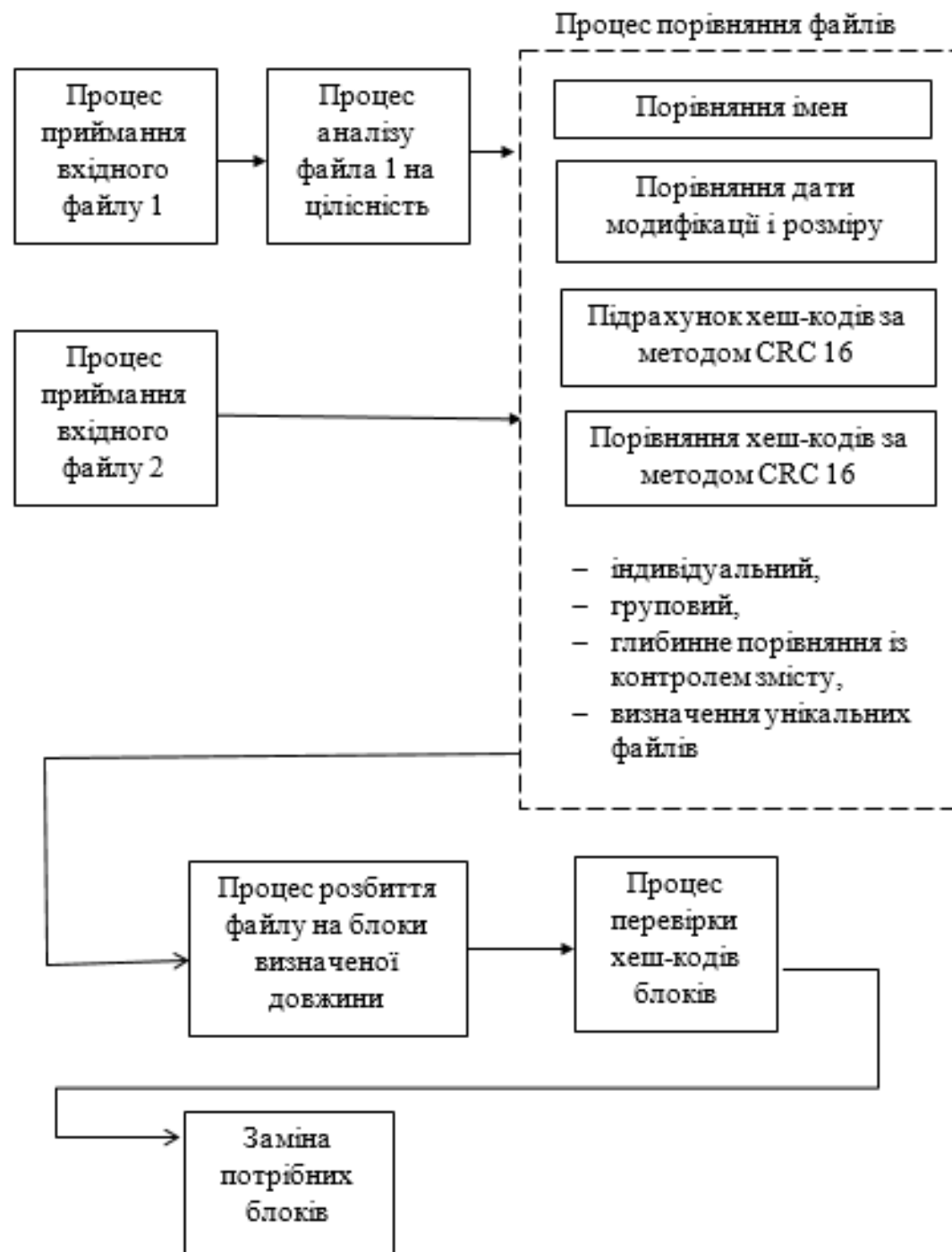
Причому працювати цей процес буде з однаковою швидкістю для файлів будь-якого розміру, адже хеш-код має фіксовану довжину.

Це дозволяє інтегрувати в програму кілька варіантів порівняльних функцій: індивідуальну, групову, глибинне порівняння із контролем змісту, а також визначення унікальних файлів. Це зробить програму зручним інструментом програміста та керівника великими спільними проектами.

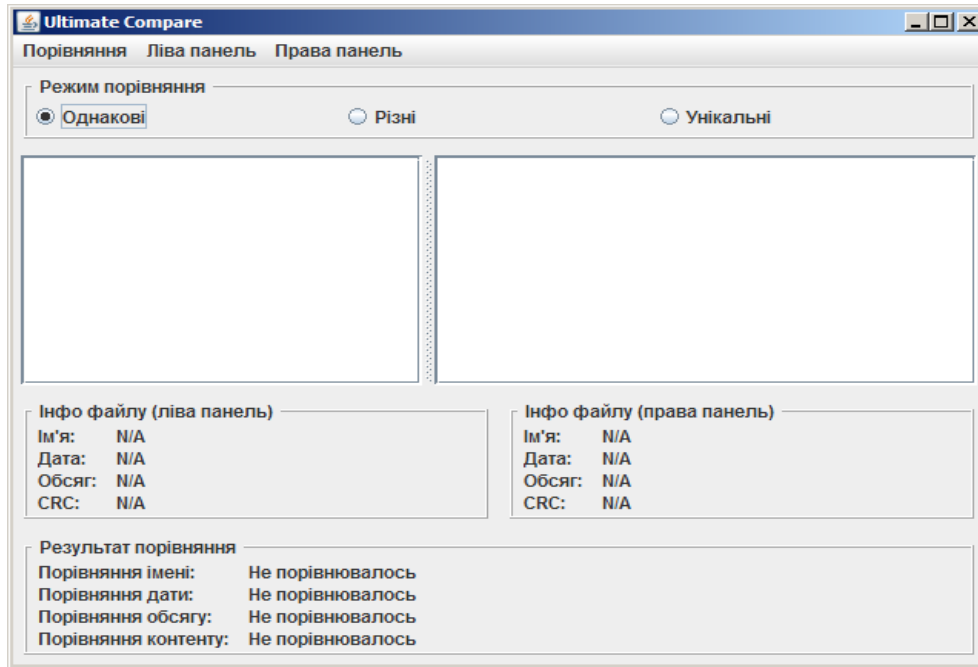
# Алгоритм роботи процесів підрахунку та порівняння хеш-кодів за методом CRC 16



# Метод блочного порівняння файлів

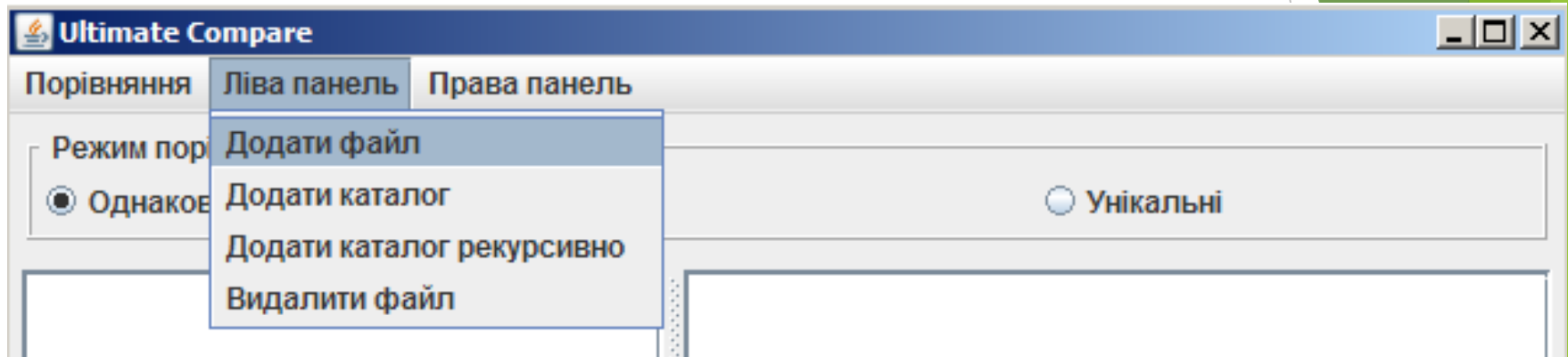


# Предмет роботи програми

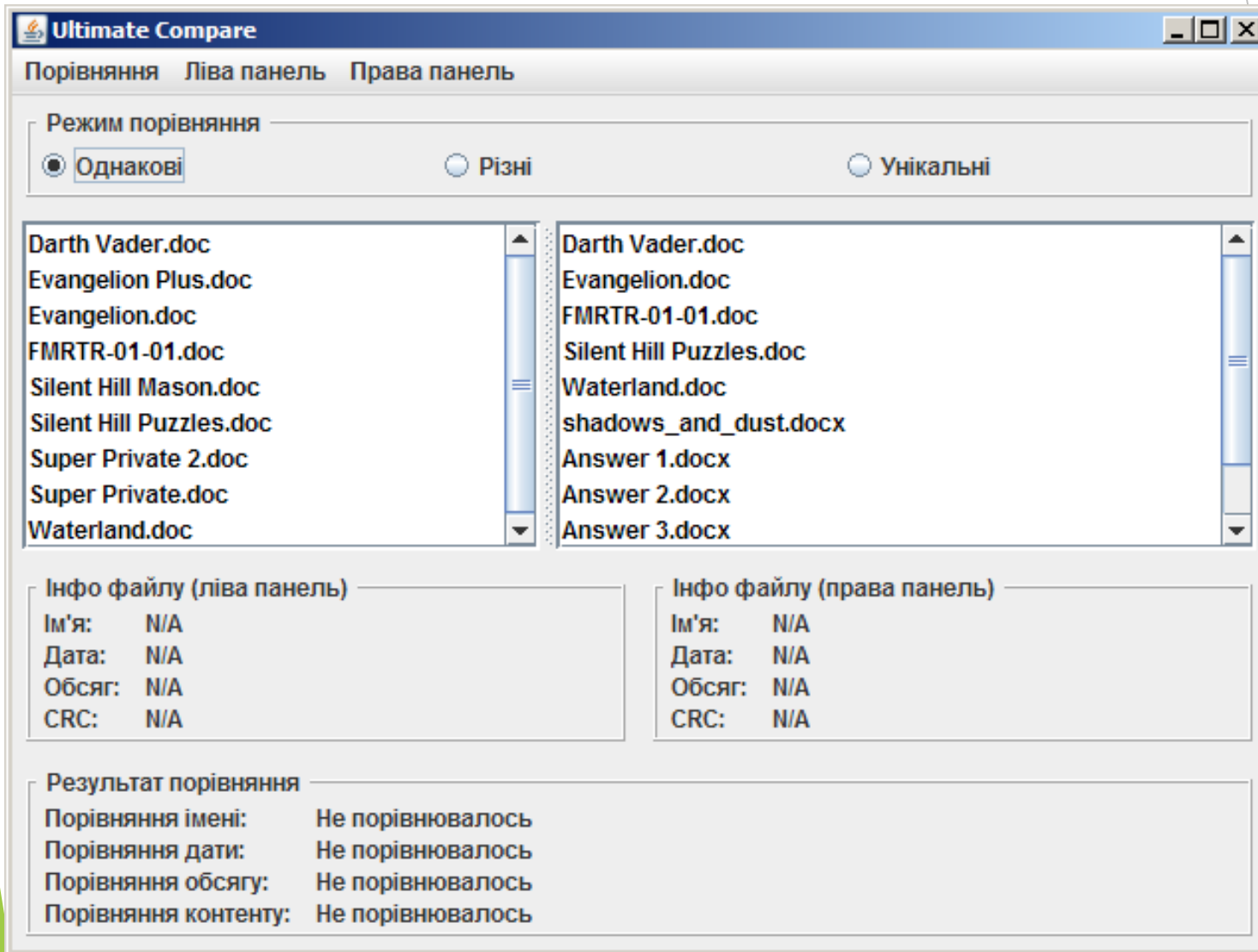


- Предметом даної роботи є програмна реалізація порівняння файлів та каталогів за допомогою хеш-коду за методом CRC 16, який дозволяє відстежувати зміни змісту файлу без витрат на рядкове чи двійкове порівняння. Саме такий метод дозволяє виконувати швидке порівняння достатньо великих обсягів даних.

# Додавання файлів на панель



# Робочий простір програми



# Порівняння різних файлів

The screenshot shows the 'Ultimate Compare' application window. The title bar reads 'Ultimate Compare'. Below the title bar, there are tabs for 'Порівняння', 'Ліва панель', and 'Права панель'. The 'Порівняння' tab is active. Underneath, there are three radio buttons for comparison modes: 'Однакові' (selected), 'Різні', and 'Унікальні'. The main area is split into two panels. The left panel lists files, with 'Silent Hill Mason.doc' selected. The right panel lists files, with 'Waterland.doc' selected. Below each list is a 'Інфо файлу' section. The left info section shows details for 'Silent Hill Mason.doc' (Date: 05.03.2008 22:43:18, Size: 38912, CRC: f5b12b4891ed15877fd4e2f54bf94eba). The right info section shows details for 'Waterland.doc' (Date: 01.08.2013 13:30:58, Size: 72192, CRC: 75921149b2de6a1ddd1ca50ccccfde3). At the bottom, the 'Результат порівняння' section shows four comparison results, all marked as 'Не співпадає' (Do not match): 'Порівняння імені', 'Порівняння дати', 'Порівняння обсягу', and 'Порівняння контенту'.

Ultimate Compare

Порівняння Ліва панель Права панель

Режим порівняння

Однакові  Різні  Унікальні

Darth Vader.doc  
Evangelion Plus.doc  
Evangelion.doc  
FMRTR-01-01.doc  
Silent Hill Mason.doc  
Silent Hill Puzzles.doc  
Super Private 2.doc  
Super Private.doc  
Waterland.doc

Darth Vader.doc  
Evangelion.doc  
FMRTR-01-01.doc  
Silent Hill Puzzles.doc  
Waterland.doc  
shadows\_and\_dust.docx  
Answer 1.docx  
Answer 2.docx  
Answer 3.docx

Інфо файлу (ліва панель)

Ім'я: Silent Hill Mason.doc  
Дата: 05.03.2008 22:43:18  
Обсяг: 38912  
CRC: f5b12b4891ed15877fd4e2f54bf94eba

Інфо файлу (права панель)

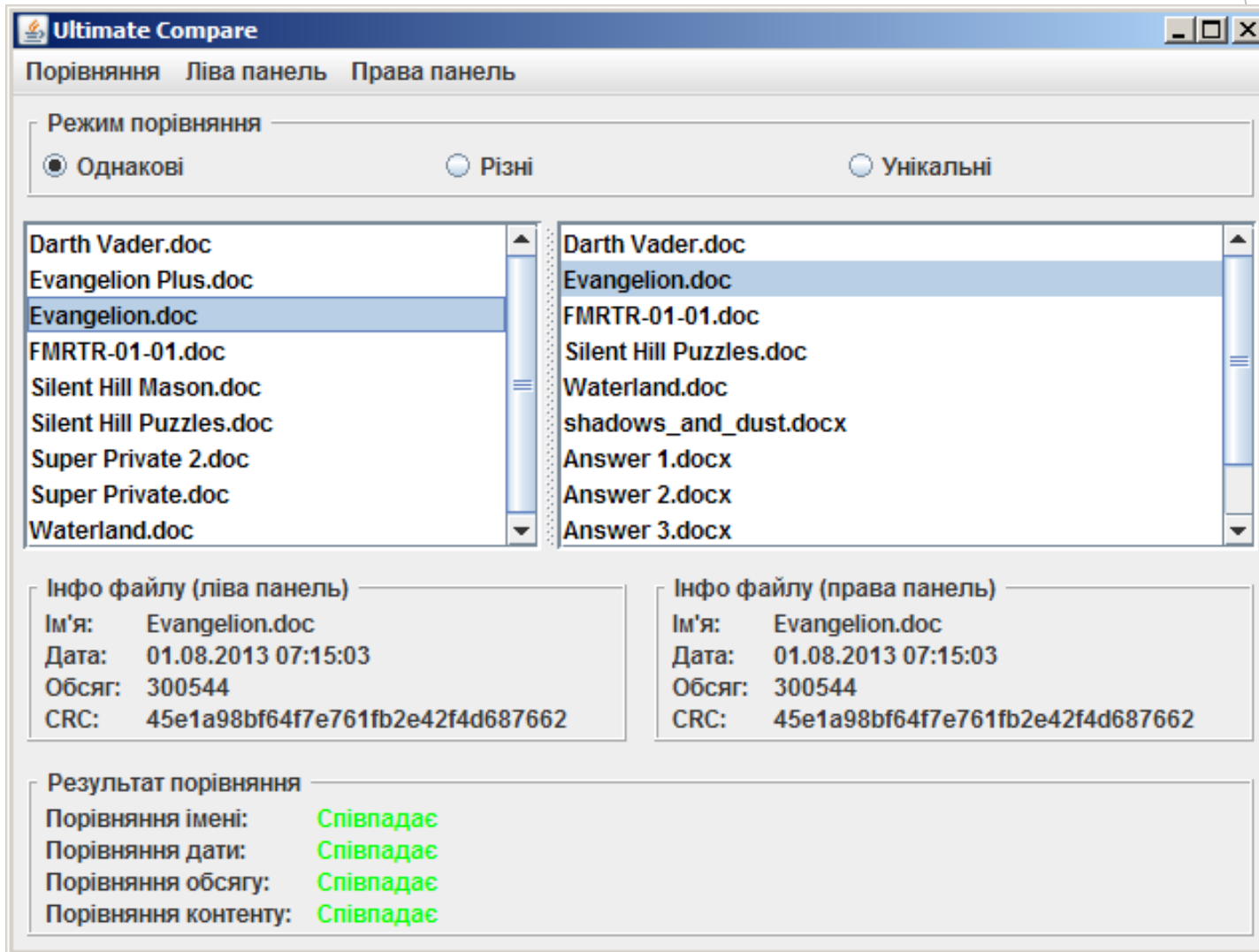
Ім'я: Waterland.doc  
Дата: 01.08.2013 13:30:58  
Обсяг: 72192  
CRC: 75921149b2de6a1ddd1ca50ccccfde3

Результат порівняння

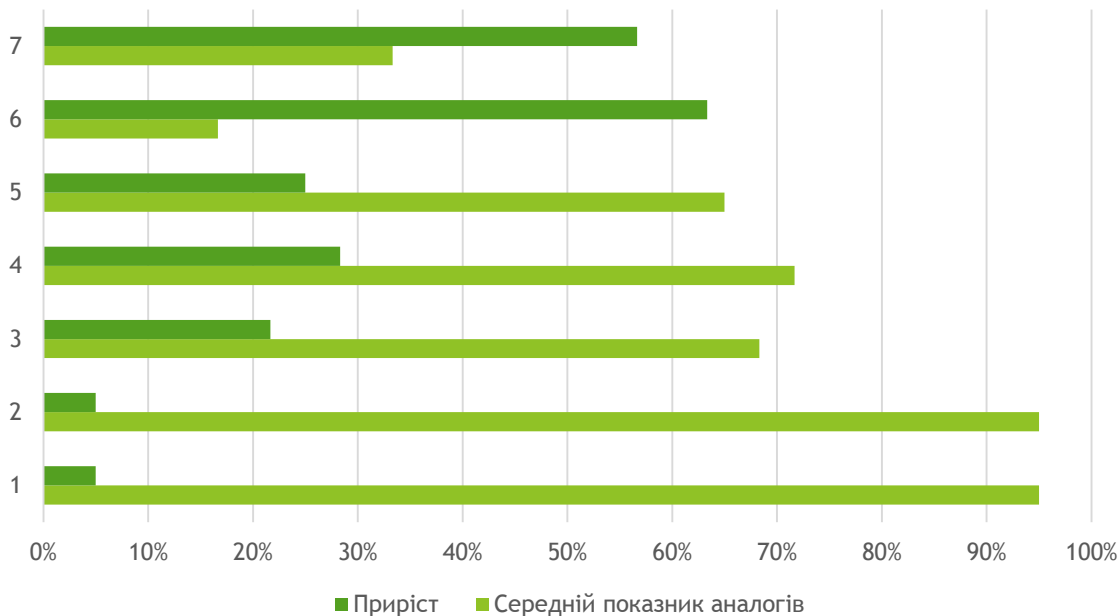
Порівняння імені: **Не співпадає**  
Порівняння дати: **Не співпадає**  
Порівняння обсягу: **Не співпадає**  
Порівняння контенту: **Не співпадає**



# Порівняння однакових файлів



# Порівняльний приріст по кожному з показників



У загальному, по ключовій позиції «Застосування хеш-даних для операції порівняння файлів» маємо приріст швидкодії під час роботи даної програми на 23%.

# ВИСНОВОК

- ▶ У результаті виконання даної дипломної роботи розроблено програмний засіб та метод блочного порівняння файлів на персональному комп'ютері по CRC-кодах засобами Java. Високий рівень вирішення поставленої задачі досягнуто за рахунок використання сучасної мови програмування Java.
- ▶ В магістерській роботі виконано дослідження і аналіз сучасних методів та засобів порівняння та синхронізації файлів, розглянуто існуючі методи вирішення задачі порівняння файлів, технологічний ланцюжок роботи методу блочного порівняння файлів та розроблено новий метод блочного порівняння файлів, запропоновано ключові процеси роботи методу блочного порівняння файлів, алгоритм роботи процесів підрахунку та порівняння хеш-кодів за методом CRC 16, який дозволяє відстежувати зміни у вмісті файлу.
- ▶ У загальному, по ключовій позиції «Застосування хеш-даних для операції порівняння файлів» маємо приріст швидкодії під час роботи даної програми на 23%.

Дякую за увагу!  
Доповідь закінчено.