

ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ МЕТОДІВ ПОШУКУ ФАЙЛІВ НА ЖОРСТКОМУ ДИСКУ

Вінницький національний технічний університет

Анотація

У даній роботі проводиться дослідження методів пошуку файлів на жорсткому диску, визначаються їх основні переваги та недоліки. Проводиться дослідження швидкісних характеристик пошуку файлів з використанням різних методів.

Ключові слова: програмний додаток, пошук, жорсткий диск, рекурсивний метод, індексний метод.

Abstract

This article research is carried out on the methods of searching files on the hard disk, determining their main advantages and disadvantages. The study of the speed characteristics of the search for files using different methods is carried out.

Keywords: program application, search, hard disk, recursive method, index method.

Вступ

Пошук файлів на жорсткому диску стає в нагоді, коли файли на комп'ютері не впорядковані по папках в достатній мірі або їх дуже багато. Для вирішення цієї проблеми використовують пошукові програми [1], кожна з яких застосовує свій метод пошуку даних на жорсткому диску, а саме рекурсивний пошук [2] або індексний пошук [3]. У роботі [4] було запропоновано новий метод пошуку файлів на жорсткому диску, використовуючи інструменти для довготривалого зберігання даних. Кожен з методів має свої переваги та недоліки, тому актуальною задачею є дослідження ефективності різних методів пошуку файлів на жорсткому диску.

Порівняння методів пошуку файлів на жорсткому диску

Файлова система в ОС «Windows» [5] має вигляд незбалансованого дерева. Традиційний рекурсивний пошук починає свою роботу з батьківського каталогу. Далі пошукова програма заходить у кожен каталог і перевіряє усі файли на відповідність заданій умові. Пошук завершується, коли програма дійшла до директорії баз дочірніх каталогів.

Час пошуку пропорційно залежить від кількості каталогів у файлової системі. Тому, чим більше каталогів, тим довше буде тривати пошук. У зв'язку з цим традиційний рекурсивний пошук файлів є досить повільним, тому доцільною є розробка нового методу пошуку даних на жорсткому диску, який підвищує швидкісні характеристики рекурсивного пошуку.

Індексний метод пошуку даних – метод пошуку, який використовує структуру даних, що містить інформацію про документи. Індексвання здійснюється пошуковою машиною та включає процес збору, сортування та зберігання даних з метою забезпечити швидкий і точний пошук інформації. Індекс – це набір файлів, в яких міститься різноманітна інформація про файли і документи, що зберігаються на диску. В індекс потрапляють різні властивості файлів: шлях, назва, розмір.

Основною перевагою даного методу над традиційним рекурсивним пошуком є те, що він значно прискорює процедуру пошуку. Недоліком індексного методу пошуку даних на жорсткому диску є те, що він зберігає у собі багато характеристик файлу, які, здебільшого, не використовуються. Це займає додаткову пам'ять та збільшує час пошуку. У зв'язку з цим, було прийнято рішення, що необхідно розробити новий метод пошуку, який буде ефективно використовувати переваги рекурсивного та індексного методів пошуку та максимально усунути їх недоліки.

В основу нового методу пошуку [4] було покладено використання інструментів для довготривалого зберігання даних на жорсткому диску: бази даних, xml-файли, текстові файли тощо. Вибір середовища зберігання даних залежить від кількості файлів у системі та необхідного ступеня захищеності інформації.

На рисунку 1 схематично відображений процес роботи нового методу пошуку з використанням інструментів для довготривалого зберігання даних (у даному випадку база даних).

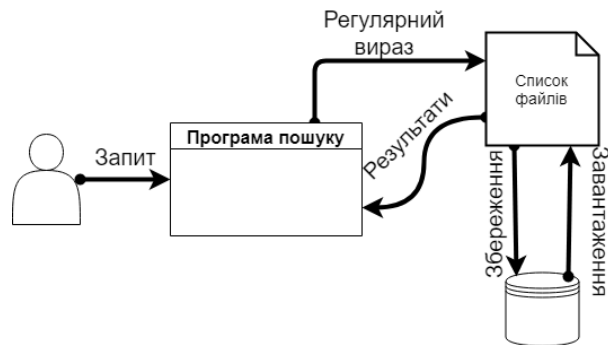


Рисунок 1 – Схема пошуку даних з використанням БД

Це працює за наступним алгоритмом:

1. Запуск пошукової програми користувачем ПК.
2. Отримання даних про файли в системі від бази даних і завантаження їх до інструментів тимчасового зберігання даних (список чи будь-який інший абстрактний тип даних).
3. Користувач подає запит у пошукову систему, а вона обробляє цей запит і здійснює пошук у списку файлів.
4. Алгоритми обробки запиту працюють уже зі списком файлів, а не із файловою системою, після цього повертають користувачеві результати пошуку.

Основною перевагою даного методу є підвищення швидкості пошуку без втрати якісних показників. Недоліком є те, що необхідно додаткове середовище для довготривалого збереження даних. Обновлення бази даних про файли може відбуватися за потребою користувача або за запланованою задачею [6] в ОС Windows.

Перевірка швидкості роботи методів пошуку файлів на жорсткому диску проводилась на наступних кількостях файлів: 1000, 10000, 100000 та 1000000. Практичні результати швидкості пошуку відображені на рисунку 2.

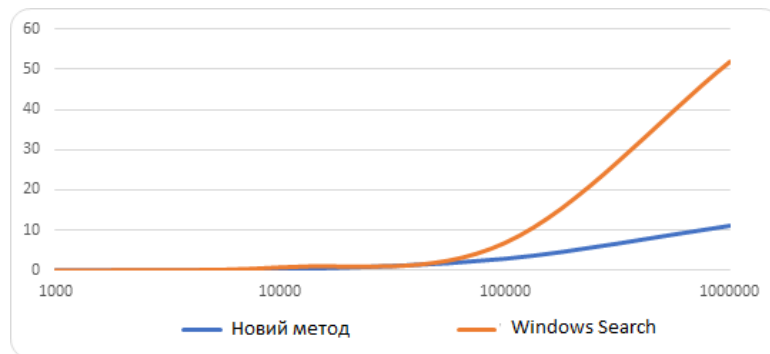


Рисунок 2 – Графік порівняння швидкості пошуку відносно кількості файлів для нового методу пошуку та вбудованого пошуку в ОС Windows

Вісь «х» відображає кількість файлів, вісь «у» час, який був затрачений на пошук файлу. Як видно з графіку на рисунку 2, при кількості файлів до 10000 обидва пошукові методи знаходять файл за секунду. Щоб знайти файл у папці зі 100000 файлів, стандартному методу пошуку в ОС Windows знадобилось 7 секунд, на відміну від нового методу, якому знадобилось 2,93 секунди. Тобто маємо прискорення процесу пошуку на 239%. Пошук файлу серед 1000000 зайняв у програми «WindowsSearch» 52 секунди, а у нового методу – 11сек., що рівне прискоренню майже в 5 разів.

Цей графік не відображає повної об'єктивності результатів порівняння швидкості пошуку двох пошукових методів. Але він показує, що новий метод пошуку даних на жорсткому диску з використанням інструментів для довготривалого зберігання даних, більш ефективний при великій кількості файлів на жорсткому диску. Тому, на даному етапі розвитку комп'ютерних комплектуючих, стає доцільно використовувати даний метод пошуку файлів на жорсткому диску.

Висновки

Отже, було досліджено основні та загальновідомі методи пошуку файлів на жорсткому диску. Було проаналізовано принципи роботи індексного та рекурсивного методів пошуку, їх основні переваги та недоліки. Після цього, було прийнято рішення, що необхідно розробити новий метод пошуку, який

буде ефективно використовувати переваги рекурсивного та індексного методів пошуку та максимально усувати їх недоліки. Даний метод використовує інструменти для довготривалого зберігання даних і підвищує швидкісні показники найпопулярніших методів пошуку даних без втрати якості пошуку до 5 разів при кількості файлів на жорсткому диску близько 1 млн.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Гаврилюк Р.О. Програмний додаток для розширеного пошуку даних на жорсткому диску «FREE SEARCHER» [Електронний ресурс] / Р.О. Гаврилюк, О.В.Романюк // Матеріали XLVI науково-технічної конференції підрозділів ВНТУ, Вінниця, 22-24 березня 2017 р. – Режим доступу: <https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/17317/3214.pdf?sequence=3>
2. Рекурсивный поиск файлов с использованием WinAPI– [Електронний ресурс]. – Режим доступу: <http://eax.me/winapi-file-search/>
3. Что такое поисковый индекс? [Електронний ресурс]. – Режим доступу: <https://www.kv.by/archive/index2009451104.htm>
4. Гаврилюк Р. О. Розробка методу пошуку файлів на жорсткому диску, використовуючи інструменти для довготривалого зберігання даних [Електронний ресурс] / Р.О. Гаврилюк, О.В. Романюк // Матеріали Міжнародної науково-практичної Інтернет-конференції «Молодь в технічних науках: дослідження, проблеми, перспективи», Вінниця, 12-17 червня 2017 року. – Режим доступу: <http://conf.inmad.vntu.edu.ua/fm/index.php?page=materials&line=29&mat=354>
5. Microsoft Windows и файловые системы – [Електронний ресурс]. – Режим доступу: <http://compress.ru/article.aspx?id=11176>
6. Планувальник завдань Windows для початківців – [Електронний ресурс]. – Режим доступу: <http://remontka.pro/windows-task-scheduler/>

Гаврилюк Роман Олександрович — студент групи ІІІ-17м, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: roman.gavriliuk@gmail.com

Романюк Оксана Володимирівна — канд. техн. наук, доцент кафедри програмного забезпечення, Вінницький національний технічний університет, Вінниця, e-mail: romaniukoksanav@gmail.com

Roman O. Gavryliuk – Student of Department of Information Technology and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: roman.gavriliuk@gmail.com

Oksana V. Romaniuk – Cand. Sc. (Eng), Assistant Professor of the Software Chair, Vinnytsia National Technical University, Vinnytsia, e-mail: romaniukoksanav@gmail.com