УДК 003.26

Г.Л. Лисенко[1]
Л.В. Кузьменко[1]

ШВИДКОДІЮЧІ ОБЧИСЛЮВАЧІ ХЕШ-ФУНКЦІЙ ДЛЯ ВЕЛИКОРОЗМІРНИХ МАСИВІВ ДАНИХ

[1] Вінницький національний технічний університет;

Анотація

Запропоновано метод підвищення швидкодії обчислення хеш-функцій у великорозмірних масивах даних для використання в блокчейн технологіях.

Ключові слова: хеш-функція, масив, матриця, алгоритм, швидкодіючі обчислювачі, логічні функції, захист інформації

Abstract

A method of increasing the speed of calculating hash functions in large data sets for use in blockchain technologies is proposed.

Keywords: hash function, array, matrix, algorithm, high-speed computers, logical functions, information security.

Вступ

Стрімкий розвиток цифрових технологій на тлі глобалізації економіки послужило основою для цифрової революції і трансформації ролі інформації з допоміжного в основний ресурс діяльності суб'єктів ринку. Перехід до цифрової економіці знаходить вияв у цифровізації бізнес-процесів, впровадженні цифрових технологій в діяльність промислових підприємств, організацій сфери послуг, державних органів, фінансових установ Незважаючи на активне освоєння цифрових технологій у всіх галузях господарської діяльності, їх можливості, переваги і недоліки вивчені ще недостатньо. Дані та інші фактори відображають актуальність теми дослідження.[1]

Метою роботи є підвищення швидкодії обчислення хеш-функцій у великорозмірних масивах даних для використання в блокчейн технологіях.

Результати дослідження

Хеш-функція – це деяка функція h(K), яка бере якийсь ключ K і повертає адресу, за якою проводиться пошук в хеш-таблиці, щоб отримати інформацію, пов'язану з K. Наприклад, K – це номер телефону абонента, а шукана інформація – його ім'я. Функція в даному випадку нам точно визначить, за якою адресою знайти шукане. Якісна хеш-функція повинна задовольняти двом вимогам: • її обчислення повинно виконуватися дуже швидко; • вона повинна мінімізувати кількість колізій. Отже, перша властивість якісної хеш-функції залежить від комп'ютера, а друга – від даних. [2].

Виокремлюють такі основні методи хешування:

1) Метод поділу. При цьому методі використовується залишок від ділення на M [2]: h(K) = K mod M. Треба ретельно вибирати цю константу. При парній константі значення функції буде парним при парному K і непарним – при непарному, що призведе до небажаного результату. Також M не повинно бути кратним трьом, оскільки при літерних ключах два з них, що відрізняються тільки перестановкою літер, можуть давати числові значення з різницею, кратній трьом.

2) Метод множення (мультиплікативний). Для цього методу хешування [2] використовується формула: h(K) = [M * ((C * K) mod 1)]

3) Мультиплікативний метод добре використовує те, що реальні файли невипадкові. Наприклад, часто множини ключів є арифметичними прогресіями, коли в файлі містяться ключі {K, K + d, K + 2d, ..., K + td}. Окремим випадком вибору константи є значення величини золотого перетину φ = (√5 – 1) / 2 ≈ ≈ +0,6180339887. Якщо взяти послідовність {φ}, {2φ}, {3φ},... де оператор {} повертає дробову частина аргументу, то на відрізку [0..1] вона буде розподілена дуже рівномірно.

4) Динамічне хешування. Існує техніка, що дозволяє динамічно змінювати розмір хешструктури.

5) Розширюване хешування. Цей метод також передбачає зміну розмірів блоків зі зростанням бази

даних, але це компенсується оптимальним використанням місця.

6) Метод ланцюжків. У разі, коли елемент таблиці з індексом, який повернула хеш-функція, вже зайнятий, до нього приєднується зв'язний список. Таким чином, якщо для кількох різних значень ключа повертається однакове значення хешфункції, то за цією адресою знаходиться показник на зв'язаний список, який містить всі значення.[3].

7) Відкрита адресація. Полягає у тому, щоб повністю відмовитися від посилань, просто переглядаючи різні записи таблиці по порядку до тих пір, поки не буде знайдений ключ K або порожня позиція.

8) Лінійна адресація використовує циклічну послідовність перевірок і описується наступним алгоритмом [1]: h(K), h(K – 1), ..., 0, M – 1, M – 2,..., h(K) + 1 Він виконує пошук ключа K в таблиці з M елементів.

9) Квадратична і довільна адресація. Замість постійної зміни на одиницю, як у випадку з лінійною адресацією, можна скористатися наступною формулою [4]: h = h + a2 , де a – це номер спроби.

10) Адресація з подвійним хешування. Цей алгоритм перевіряє таблицю трохи інакше, тримаючи її обома хеш-функції h1(K) і h2(K). Остання повинна породжувати значення в інтервалі від 1 до M – 1, взаємно прості з M.

11) Видалення елементів хеш-таблиці. Взагалі кажучи, обробляти видалення можна, позначаючи елемент як видалений, а не як порожній. Таким чином, кожна клітинка в таблиці буде містити вже одне з трьох значень: порожня, зайнята, видалена. При пошуку вилучені елементи будуть трактуватися як зайняті, а при вставці – як порожні, відповідно.

12) Застосування хешування. Одне з побічних застосувань хешування полягає в тому, що воно створює свого роду зліпок, «відбиток пальця» для повідомлення, текстового рядка, області пам'яті і т. п. Такий «відбиток пальця» може прагнути як до «унікальності», так і до «схожості». В цій якості однією з найважливіших областей застосування є криптографія.

13) Хешування паролів. Хешування паролів – метод, що дозволяє користувачам запам'ятовуються не 128 байт, тобто 256 шістнадцяткових цифр ключа, а деякий осмислений вираз, слово або послідовність символів, що називається паролем. [4].

Для обробки великорозмірних масивів даних слід використовувати алгоритми хешування функцій і яких довжина хеша не менше ніж 256 біт. Такі алгоритми зображені в таблиці 1.

Таблиця 1 – Алгоритми хешування даних

Назва	**Довжина хеша (біти)**	**Максимальна довжина повідомлення**
SHA2	224/256/384/512	2^{64}
SHA-3	224/256/384/512	2^{128}
WHIRLPOOL	512	2^{64}
EDON-R,	256/512	2^{64}

Хешфункцією в даному випадку називається таке математичне або алгоритмічне перетворення заданого блоку даних, яке має такі властивості: 1. Хеш-функція має нескінченну область визначення. 2. Хеш-функція має кінцеву область значень. 3. Вона необоротна. 4. Зміна вхідного потоку інформації на один біт змінює близько половини всіх біт вихідного потоку, тобто результату хеш-функції. Ці властивості дозволяють подавати на вхід хеш-функції паролі, тобто текстові рядки довільної довжини будь-якою національною мовою і, обмеживши область значень діапазоном 0..2N – 1, де N – довжина ключа в бітах, отримувати на виході досить рівномірно розподілені по області значення блоки інформації – ключі. [5]

Висновки

Встановлено, що запропонований підхід дозволяє підвищити швидкість обробки великорозмірних масивів даних. Також в цій роботі розглянуто найбільш відомі алгоритми хешування із певною кількістю довжини хеша, також описані основні методи хешування даних. Вибрано основні типи алгоритму для конкретної практичної задачі яка поставлена в меті роботи.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Шнайер, Брюс. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си / Брюс Шнайер. – М.: Триумф, 2002. – ISBN 5-89392-055-4.

2. Кнут Дональд. Искусство программирования. Том 3. Сортировка и поиск = The Art of Computer Programming, vol. 3. Sorting and Searching / Дональд Кнут. – 2-е издание. – М.: Вильямс, 2007. – С. 824. – ISBN 0-201- 89685-0.

3. Кормен Т. Алгоритмы: построение и анализ. / Т. Кормен, Ч. Лейзерсон, Р. Ривест. – М.: МЦНМО, 2001.

4. Вирт Никлаус. Алгоритмы и структуры данных / Никлаус Вирт. – М.: Мир, 1989. – ISBN 5-03-001045-9.

5. Lysenko, G. L., Kuzmenko, L. V., Kisała, P., Klimek, J., & Kalimoldayev, M. The use of optically controlled transparent and blockchain technology for the processing of large-scale data arrays. In: Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019. International Society for Optics and Photonics, 2019. p. 111760G.

Кузьменко Лілія Вікторівна – аспірантка кафедри Лазерної та оптикоелектронної техніки, Вінницький національний технічний університет, м. Вінниця, Україна

Науковий керівник: ***Лисенко Геннадій Леонідович*** – к.т.н., проф. кафедри лазерної та оптикоелектронної техніки, Вінницький національний технічний університет, м. Вінниця, Україна.

Kuzmenko Liliia V – Phd student of the Department of Laser and Optoelectronic Technology, Vinnytsia National Technical University, Vinnytsia, Ukraine.

Supervisor: ***Lysenko Gennadii L*** - candidate of technical sciences, prof. Department of Laser and Optoelectronic Technology, Vinnytsia National Technical University, Vinnytsia, Ukraine.