

АНАЛІЗ АТАК НА ДЕРЕВОПОДІБНІ ГЕШ-ФУНКЦІЇ НА ОСНОВІ МУЛЬТИКОЛІЗІЙ

Вінницький національний технічний університет

Анотація

Розглянуто атаки на деревоподібні геш-функції, що базуються на мультиколізіях. Досліджено можливі методи компрометації функцій гешування шляхом генерації множинних колізій, а також їхній вплив на криптографічну стійкість деревоподібних структур. Проведено аналіз ефективності відомих контрзаходів для захисту деревоподібних геш-функцій від атак мультиколізій. Результати дослідження можуть бути використані для вдосконалення механізмів безпеки сучасних криптографічних алгоритмів.

Ключові слова: геш-функції, деревоподібні геш-функції, атаки на геш-функції, мультиколізії, криптографічна стійкість, гешування.

Abstract

The attacks on tree-based hash functions based on multicollisions are considered. Possible methods of compromising hashing functions through the generation of multiple collisions and their impact on the cryptographic resistance of tree structures are explored. An analysis of the effectiveness of existing countermeasures to protect tree-based hash functions from multicollision attacks has been conducted. The research results can be used to improve the security mechanisms of modern cryptographic algorithms.

Keywords: hash functions, tree-based hash functions, hash function attacks, multicollisions, cryptographic resistance, hashing.

Вступ

Деревоподібні геш-функції широко використовуються в криптографічних застосунках, таких як блокчейн, цифрові підписи та системи автентифікації. Проте їхня безпека залежить від стійкості до атак, заснованих на колізіях. Однією з найсерйозніших загроз є атаки на основі мультиколізій, що дозволяють зловмисникам компрометувати криптографічну цілісність системи. У цьому дослідженні проаналізовано вразливості деревоподібних геш-функцій до таких атак та запропоновано методи їхньої нейтралізації.

Метою даного дослідження є підвищення криптографічної стійкості геш-функцій до атак на основі мультиколізій.

Для досягнення мети було розв'язано такі завдання:

- виявити основні вразливості деревоподібних геш-функцій до атак мультиколізійного типу;
- оцінити ефективність відомих механізмів захисту від мультиколізій та їхній вплив на криптографічну безпеку;
- запропонувати можливі методи підвищення стійкості деревоподібних геш-функцій до атак.

Мультиколізійні атаки на деревоподібні геш-функції

Атаки мультиколізійного типу є однією з ключових загроз для сучасних криптографічних систем, що використовують геш-функції. Вони засновані на знаходженні множинних колізій, що дозволяє значно знижувати криптографічну стійкість алгоритмів гешування, зокрема тих, що реалізовані на основі каскадної конструкції Меркля-Дамгаарда.

Концепція мультиколізії вперше була запропонована Антуаном Жу, який показав, що якщо для певної геш-функції можливо знайти одиничну колізію зі складністю $O\left(2^{\frac{2n}{3}}\right)$, то складність множинних колізій зростає лінійно в той час як їх кількість — експоненційно [1]. Це досягається шляхом рекурсивного розширення множинності колізій: спочатку знаходиться одна колізійна пара, потім на її основі будується ще одна, після чого вони комбінуються так, що кількість різних повідомлень, які мають однаковий геш, подвоюється на кожному етапі. Після декількох ітерацій цього процесу зловмисник отримує набір із різних повідомлень, усі з яких мають однаковий геш.

Особливість атаки Жу полягає в тому, що її реалізація можлива для будь-якої геш-функції, що побудована за принципом послідовного обчислення блоків даних. Більшість традиційних криптографічних геш-функцій, таких як MD5, SHA-1 і навіть SHA-256, використовують саме такий підхід, що робить їх потенційно вразливими до подібних атак. Проблема полягає в тому, що будь-яка зміна блоку на початкових рівнях обчислення автоматично впливає на всі наступні рівні, проте при наявності мультиколізійного механізму можна знаходити такі варіанти вхідних даних, які дадуть той самий результат гешування [1].

Значна небезпека атаки мультиколізійного типу полягає в тому, що вона може бути використана для компрометації цифрових підписів та блокчейн-систем. Наприклад, коли користувач підписує певний документ електронним цифровим підписом. Якщо зловмисник володіє механізмом побудови мультиколізій, він може підготувати декілька варіантів повідомлення таким чином, що всі вони матимуть однаковий геш. Це означає, що після підписання одного з таких документів можна легко підмінити його на будь-який інший із того ж набору, і підпис залишатиметься дійсним. Аналогічна загроза існує для блокчейн-систем, де транзакції організовані у вигляді дерева Меркла. Якщо знайдено мультиколізію для певного вузла цього дерева, то можливо замінити групу транзакцій на іншу без зміни підсумкового кореневого гешу.

Окрім класичної мультиколізії, до основних видів атак мультиколізійного типу можна віднести [2]:

1) Прогресивна мультиколізія – варіант класичної атаки, при якому використовується рекурсивний підхід для побудови множинних колізій у багаторівневих структурах. Це дозволяє атакувати деревоподібні геш-функції шляхом вибору вузлів із повторюваними гешами та формування множини варіантів з однаковим підсумковим значенням.

2) Атаки на основі розширення повідомлення – застосовуються в геш-функціях із деревоподібною структурою, де гешування здійснюється блоками. Використовуючи знання про внутрішній стан геш-функції, зловмисник може генерувати нові блоки, які дадуть той самий підсумковий геш. Це особливо небезпечно в системах, що використовують дерева Меркла або інші багаторівневі схеми.

3) Атаки на зменшення ентропії гешу – у разі використання геш-функцій зі слабкими або ущільненими підсумковими значеннями зловмисники можуть знаходити множинні комбінації вхідних даних, що приводять до однакового гешу. В деревоподібних структурах це призводить до можливості маніпулювати окремими гілками, зберігаючи правильний геш кореневого вузла.

Атаки мультиколізійного типу становлять серйозну загрозу для деревоподібних геш-функцій, основні вразливості таких структур включають залежність від дерева, можливість маніпуляції окремими вузлами, низьку ентропію на проміжних рівнях.

Відомі механізми захисту від мультиколізій

Одним із ключових методів захисту деревоподібних геш-функцій є впровадження механізму динамічного гешування вузлів, що передбачає додавання випадкових значень (salt) до кожного вузла дерева перед обчисленням його гешу. Такий підхід унеможливує підбір мультиколізійних комбінацій для окремих гілок дерева, оскільки навіть ідентичні вхідні дані на різних рівнях дерева будуть генерувати різні геш-значення. Традиційні схеми деревоподібного гешування використовують жорстко детермінований порядок обчислення гешів вузлів, що дозволяє зловмиснику знаходити колізії для окремих піддерев та переносити їх угору по структурі [3]. Введення випадкових доповнень порушує цю передбачуваність, роблячи неможливим знаходження колізій для всіх вузлів одночасно. Практична реалізація динамічного гешування може здійснюватися кількома способами:

- статичне випадкове доповнення, де при генерації дерева кожному вузлу додається унікальне випадкове значення, що зберігається разом із вузлом і використовується під час перевірки цілісності;
- динамічне випадкове доповнення, де випадкові значення обчислюються при кожному повторному гешуванні дерева, що забезпечує додатковий рівень захисту від зворотного аналізу зловмисником.

Ще одним важливим механізмом захисту деревоподібних геш-функцій є застосування динамічної перестановки гілок перед обчисленням гешу вузла. У класичних схемах дерева порядок гешування вузлів є фіксованим, що дає змогу зловмиснику попередньо аналізувати структуру та будувати атаки з підміною окремих гілок дерева [3]. Перестановка гілок усуває цю передбачуваність, оскільки порядок обчислення гешу може змінюватися випадковим чином на кожному рівні структури. Основні варіанти реалізації цього методу включають:

- випадковий порядок гешування вузлів, де на кожному рівні випадковим чином змінюється порядок обчислення геш-значень дочірніх вузлів;

- закодована перестановка на основі атрибутів вузлів, де порядок гешування може залежати від певних унікальних параметрів вузлів, наприклад, їхніх цифрових ідентифікаторів або зовнішніх параметрів системи.

Також значну роль у запобіганні атакам мультиколізійного типу відіграє контроль структури вхідних даних перед їхнім гешуванням. Навіть при використанні випадкових доповнень і перестановки гілок залишається загроза маніпуляцій із підміною частин вхідних даних так, щоб сформувати потрібний підсумковий геш. Щоб запобігти цьому, можуть застосовуватися наступні механізми [4]:

- перевірка унікальності вхідних даних перед додаванням у структуру, де система перевіряє, чи не містять вхідні дані прихованих дублікатів або структурних особливостей, що можуть бути використані для атаки;

- додавання до кожного вузла контрольних параметрів, де крім геш-значення, вузол може містити унікальні мета-дані (наприклад, часові мітки, цифрові підписи), що робить неможливим підміну окремих гілок без зміни кінцевого значення;

- використання багаторівневого гешування, де додаткове гешування даних перед їхнім включенням у структуру дозволяє ускладнити пошук колізій на проміжних етапах.

У таблиці 1 представлено аналіз основних механізмів захисту деревоподібних геш-функцій від атак мультиколізійного типу, що включають динамічне гешування вузлів, перестановку гілок перед гешуванням та контроль структури вхідних даних. Зокрема, у таблиці наведено їхні позитивні та негативні сторони, що дозволяє оцінити ефективність кожного з методів у контексті забезпечення криптографічної стійкості.

Таблиця 1. Аналіз основних механізмів захисту деревоподібних геш-функцій від атак мультиколізійного типу

Механізм	Позитивні сторони	Негативні сторони
Динамічне гешування вузлів	Підвищена стійкість	Потреба у зберіганні випадкових значень
	Захист від атак другого прообразу	Ускладнення перевірки цілісності
	Відсутність додаткового впливу на розмір гешу	Можливість атаки через повторне використання salt
Перестановка гілок перед гешуванням	Захист від предиктивних атак	Ускладнення процесу перевірки гешу
	Підвищена гнучкість структури дерева	Можливість зниження ефективності у певних реалізаціях
	Мінімальний вплив на розмір та швидкість гешування	Необхідність узгодження на рівні всіх учасників системи (у випадку блокчейна)
Контроль структури вхідних даних	Запобігання атакам із підміною даних	Збільшення розміру вузлів геш-дерева
	Підвищена прозорість обчислення гешів	Додаткові обчислювальні витрати
	Покращена стійкість до комбінованих атак	Можливість хибнопозитивних спрацьовувань

. Динамічне гешування вузлів забезпечує підвищену стійкість до атак та захищає від атак другого прообразу, однак вимагає додаткового зберігання випадкових значень і ускладнює перевірку цілісності. Перестановка гілок перед гешуванням робить систему стійкішою до предиктивних атак, проте ускладнює процес перевірки гешу та може впливати на ефективність у певних реалізаціях. Контроль структури вхідних даних запобігає атакам із підміною даних і покращує прозорість обчислення, але водночас збільшує розмір вузлів геш-дерева та потребує додаткових обчислювальних ресурсів.

Висновки

Проведене дослідження дозволило виявити основні вразливості деревоподібних геш-функцій до атак мультиколізійного типу та оцінити ефективність відомих механізмів їхнього захисту. Аналіз показав, що однією з ключових проблем таких структур є залежність від порядку гешування, що створює можливість маніпуляції окремими вузлами дерева без зміни кореневого гешу. Низька ентропія проміжних рівнів гешування також сприяє підбору мультиколізій, що може використовуватися для компрометації криптографічних систем. Найбільш небезпечними є атаки на основі розширення повідомлення, які дозволяють змінювати частини дерева без порушення його загальної структури, що особливо критично для блокчейну та цифрових підписів.

Оцінка ефективності механізмів захисту продемонструвала, що найбільш перспективними

методами є динамічне гешування вузлів, перестановка гілок перед гешуванням та контроль структури вхідних даних. Динамічне гешування забезпечує унікальність кожного вузла шляхом додавання випадкових значень, що значно ускладнює реалізацію атак мультиколізійного типу. Однак цей метод потребує додаткових ресурсів для зберігання випадкових параметрів, що може впливати на продуктивність системи. Перестановка гілок перед гешуванням дозволяє зменшити ризик предиктивних атак, ускладнюючи підбір колізій на проміжних рівнях дерева. Водночас, ця методика потребує синхронізації між вузлами, що може ускладнити її реалізацію у розподілених системах. Контроль структури вхідних даних є ефективним способом запобігання атакам на підміну та дозволяє ідентифікувати потенційні загрози на ранніх етапах гешування, проте його впровадження може значно збільшити обсяг необхідних обчислювальних ресурсів.

Підвищення криптографічної стійкості деревоподібних геш-функцій можливе шляхом комплексного використання всіх трьох розглянутих механізмів. Поєднання динамічного гешування, перестановки гілок та контролю вхідних даних дозволяє досягти високої безпеки при мінімальному впливі на продуктивність системи. Подальші дослідження у цьому напрямку можуть бути спрямовані на розробку гібридних підходів до структурного захисту геш-функцій, зокрема шляхом адаптації кванто-стійких алгоритмів.

Значення отриманих результатів полягає у можливості їх застосування для вдосконалення сучасних криптографічних протоколів, що використовують деревоподібні геш-функції. Розробка стійких до мультиколізій алгоритмів є критично важливою для забезпечення безпеки у сферах блокчейну, цифрових підписів, контролю цілісності даних та безпечних систем аутентифікації. Використання комбінованих підходів дозволяє знайти оптимальний баланс між рівнем безпеки, швидкістю обробки та ефективністю використання обчислювальних ресурсів, що сприятиме подальшому розвитку криптографічних методів захисту.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. В. А. Лужецький, Ю. В. Баришев. Конструкції хешування стійкі до мультиколізій. Наукові праці ВНТУ, 2011 р. С. 1-8. URL: <https://praci.vntu.edu.ua/index.php/praci/article/view/191/189>. (дата звернення: 10.02.2025).
2. Jan Buzek, Stefano Tessaro. Collision Resistance from Multi-collision Resistance for All Constant Parameters. *Advances in Cryptology – CRYPTO 2024*. pp. 429–458. URL: https://link.springer.com/chapter/10.1007/978-3-031-68388-6_15. (дата звернення: 10.02.2025).
3. Ron D. Rothblum, Prashant Nalini Vasudevan. Collision Resistance from Multi-collision Resistance. *Journal of Cryptology*, Volume 37, article number 14, 2024. p. 26. URL: <https://link.springer.com/article/10.1007/s00145-024-09495-5>. (дата звернення: 10.02.2025).
4. Juha Partala. Indifferentiable hash functions in the standard model. *IET Information Security*, 2021. pp. 309-316. URL: <https://ietresearch.onlinelibrary.wiley.com/doi/full/10.1049/ise2.12025>. (дата звернення: 10.02.2025).

Казміревський Віталій Віталійович — аспірант кафедри Захисту інформації, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: kazmirevskiy1999@gmail.com

Науковий керівник: **Баришев Юрій Володимирович** — к. т. н., доцент кафедри захисту інформації, Вінницький національний технічний університет, Вінниця. email: yuriy.baryshev@vntu.edu.ua

Vitalii Kazmirevskiy — postgraduate student of Information Protection Department, Faculty for Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, email: kazmirevskiy1999@gmail.com

Supervisor: **Yurii Baryshev** – PhD (Eng), Associate Professor of Information Protection Department, Vinnytsia National Technical University, Vinnytsia. email: yuriy.baryshev@vntu.edu.ua