УДК 621.3.06

АНАЛІЗ ЗАСТОСУВАННЯ ФУНЦІЇ ГЕШУВАННЯ У ТЕХНОЛОГІЇ BLOCKCHAIN

П. В. КРАВЧУК, І. Д. ГОРБЕНКО, А. І. ПУШКАРЬОВ

Наведено результати аналізу функцій гешування для їхнього застосування у системах, що використовують технологію Blockchain, а також результати порівняльного аналізу їх основних властивостей та рекомендації щодо застосування.

Ключові слова: геш-значення, електронний підпис, криптографічні механізми, криптографічна стійкість, послуги безпеки, технології блокчейн, складність криптопекретворень, функція гешування.

ВСТУП

Аналіз показав, що, під час розробки технологій Blockchain (далі – «блокчейн»), необхідно враховувати важливі аспекти та вимоги до технології блокчейн [у тому щодо інформаційної та кібербезпеки 1–4].

Одним із найважливіших криптографічних компонентів цієї технології, що суттєво визначає їх захищеність – є функції гешування відповідних даних. Тому, під час проектування технологій блокчейн геш-функція, що застосовуватиметься, має бути вибрана за основними безумовними та умовними критеріями – криптографічна стійкість проти класичних та квантових атак, складність (швидкодія) тощо. Важливо також при виборі функції гешування розглянути та порівняти основні альтернативи, у тому числі рівень їх стандартизації та тенденції застосування, а також популярність щодо застосування в сучасних мережах

Метою статті є аналіз та порівняльний аналіз функцій гешування за основними такими безумовними критеріями як криптографічна стійкість проти класичних та квантових атак, складністю криптографічних перетворень (швидкодії) та можливістю і умовами застосування в технологіях блокчейн.

1. ЗАГАЛЬНІ ПОЛОЖЕННЯ

1.1. Сутність технології блокчейн

Блокчейн (англ. Blockchain, Block chain від block – блок, chain – ланцюг) [1] – це незмінні системи цифрових регістрів, реалізовані розподіленим чином (тобто, без центрального сховища) та зазвичай без центрального органу. На самому базовому рівні вони дозволяють спільноті користувачів записувати транзакції в загальнодоступному реєстрі цієї спільноти, так що ніяка транзакція не може бути змінена після опублікування.

Важливим компонентом технології блокчейн є використання криптографічних функцій ґешування (ФГ) для багатьох операцій, перше за все, таких як гешування вмісту блоку. Ґешування [2] це метод одно направленого відображення вхідних даних (файла, даних, деякого тексту або зображення тощо) довільної довжини в унікальне вихідне значення фіксованого розміру L_h (називається просто дайджест). З великою ймовірністю вихідне значення вважається випадковим, найменша зміна уведних даних (навіть на один біт) призводить до зміни вихідного, в середньому на один біт.

1.2. Функції гешування

Функції гешування *h(x)* (ФГ) [2,3] – це однонаправлена колізійна стійка функція відображення, що приймає на вході як аргумент інформаційну послідовність (рядок) *M* довільної довжини L_m і дає на виході практично випадкову послідовність (рядок) фіксованої довжини *Lh*. Результат гешування інформаційної послідовності *M* називають геш - образом *h(M)* Для відомих функцій гешування співвідношення між довжинами *M* і *h(M)* може бути довільним, тобто

$$|M| > |h(M)|, |M| < |h(M)|, |M| = |h(M)|.$$

Оскільки результат роботи функції гешування називається геш-образом, то масив даних *M* зазвичай називають прообразом (першим прообразом).

Наведемо формальне визначення функції гешування. Нехай *{0, 1}^m* – безліч всіх двійкових рядків довжини *m, {0, 1}** – безліч всіх двійкових рядків кінцевої довжини. Тоді геш функцією *h* називається перетворення виду:

$$h: \{0, 1\}^* \rightarrow \{0, 1\}^m,$$

де *m* – розрядність геш-образу. На рис. 1 як приклад показано схему гешування, PRNG (Pseudo-Random Number Generator) – генератор псевдовипадкових чисел; Q – елементи пам'яті PRNG; h_0 – вектор ініціалізації; $n=|m_i|$ – розрядність блоків інформаційної послідовності, $i = 1,\ldots, t$:

$$M = m_1, \ldots, m_t,$$

t – число блоків послідовності *M*; *N* – число елементів пам'яті PRNG. Для PRNG процес отримання функції гешування [4] можна спрощено розглядати як накладення псевдовипадкової послідовності (PRS, Pseudo-Random Sequence) на вхідну послідовність для подальшого її перетворення.

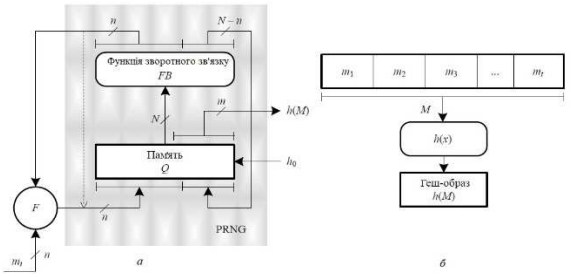

Рис. 1. Спрощений механізм функції гешування: а PRS

Однією із основних вимог до функцій гешування є їх колізійна стійкість. Фізично її можна визначити як можливість знайти колізію функції гешування, тобто складність знаходження двох різних випадкових даних (рядків) M1 і M2, таких що h(M1) = h (M2). Інакше кажучи, коли для двох різних аргументів M1 і M2 значення ФГ збігаються.

На рис. 2 приклад виникнення колізії при гешуванні даних M_3 і M_4.

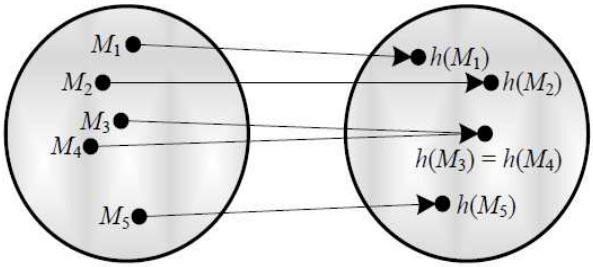

Рис. 2. Відображення безлічі прообразів і геш-образів

1.3. Вимоги до функцій гешування

ФГ знаходять широке застосування, в тому числі необхідно виділити такі застосування ФГ [2, 3,4]:

– механізми контролю цілісності та справжності інформації та ресурсів;

– протоколи електронного підпису, електронного штампу та мітки часу);

– алгоритми генерація псевдовипадкових послідовностей;

– криптографічні протоколи узгодження та встановлення ключів;

– контроль цілісності баз даних, у тому числі без використання спільного секрету тощо.

В технології блокчейн ФГ використовуються для контролю цілісності повідомлень (блоків) [5], які передаються по мережі або зберігаються поза межами захищеного середовища. Для ФГ, що використовуються у випадку технології блокчейн для контролю цілісності баз даних, мають виконуватися такі вимоги [5,2]:

– відновлюваність у просторі і часі;

– детермінованість – при однакових вхідних даних результат виконання ФГ буде однаковим (одне і те саме повідомлення завжди призводить до одного й того ж гешу);

– стійкість до знаходження колізій;

– стійкість до знаходження прообразу – неможливість знаходження невідомого прообразу для будь-яких заданих геш-значень;

– стійкість до атаки пошуку другого прообразу;

– атака збільшення довжини (length extension attack)

– атака фіксованих точок (fixed points).

Оцінки ідеальної стійкості для ФГ [2, 6] наведені на рис. 3.

Тип геш-функції	Ціль атаки	Ідеальна стійкість
Однонаправлена ФГ	Знаходження прообразу	2^l
	Знаходження другого прообразу	2^l
	Знаходження колізії	$2^{\frac{l}{2}}$
	Збільшення довжини	2^l
	Фіксовані точки	2^l
Функція створення MAC	Точне знайдення ключа	$\left\lceil \frac{k}{l} \right\rceil + \frac{2^k - 1}{1 - 2^{-l}}$
	Підробка повідомлення	$P_m = \max(2^{-k}, 2^{-l})$

Рис. 3. Необхідна стійкість ФГ до атак

1.4. Вибір перспективних ФГ

Розглянемо загальні дані щодо перспективних (сучасних) ФГ [7, 2]. У системах, що побудовані на технології блокчейн, сьогодні використовуються тільки деякі ФГ, що стандартизовані на міжнародному рівні та NIST США. Зокрема, найбільш поширені блокчейн розробки (крипто валюти), у яких використовують наступні стандартизовані алгоритми гешування: SHA-256, EtHash, Scrypt, X11, CryptoNight, EquiHash.

Оскільки більшість популярних криптовалют використовують ФГ SHA-2 та SHA-3 («Keccak») [8], то розглянемо їх порівняно з іншими відомими перспективними стандартизованими ФГ: російським ГОСТ 34.11-2012 («Стрибог»), Whirlpool та українським стандартом ДСТУ 7564:2014 («Kupyna»).

На рисунку 4 наведено порівняння ФГ за основними загальними параметрами.

Вибрані ФГ мають розмір вихідного геш-значення від 224 біт (доцільніше використовувати мінімальний вихід у 256 біт) до 512 біт. Інші показники, як розмір внутрішнього стану, розмір блоку та слова – лежать у схожих межах та залежать здебільшого від структури самого алгоритму гешування. За цими даними виділяти кращу ФГ не є правильним, адже це не говорить напряму про стійкість та швидкість алгоритму.

Алгоритм	Розмір виходу, біт	Розмір внутрішнього стану, біт	Розмір блоку, біт	Розмір слова, біт	Кількість раундів
ГОСТ 34.11-2012	256 (512)	256(512)	512	32	12
SHA-2 256	256/224	256	512	32	64
SHA-2 512	512/384	512	1,024	64	80
SHA-3 256	256	1600	1088	64	24
SHA-3 512	512	1600	576	64	24
Whirlpool	512	512	512	8	10
Kupyna 256	256	256	512	64	10
Kupyna 512	512	512	1024	64	14

Рис. 4. Загальні дані алгоритмів сучасних ФГ

2. МЕТОДИ ДОСЛІДЖЕНЬ ФГ

2.1. Статистичне тестування ФГ

Статистичне тестування ФГ, що аналізуються, із використанням методик тестування, що визначені в NIST STS 800- 22(2009) [9].

Тестування статистичних властивостей виконувалось у таких режимах:

1) висока збитковість вхідної послідовності;
2) у режимі генератора псевдовипадкових послідовностей відповідно до ISO/IEC 18031 із використанням функції гешування, з виконанням реініціалізації;
3) у режимі генератора псевдовипадкових послідовностей відповідно до ISO/IEC 18031 із використанням функції гешування, без виконання реініціалізації.

2.2. Порівняння складності (швидкодії) ФГ

Важливим критерієм для порівняння ФГ є складність (швидкодія) гешування.

Для оцінки складності реалізації ФГ вимірювались швидкість роботи стандартизованих реалізацій ФГ мовою Java. Для цього було використано компілятор IntelliJ IDEA 2016.2.4(64). Тестування швидкодії гешування проведено на комп'ютері Intel Core i5-4460 3.2 GHz, 8 GB RAM під управлінням операційної системи Microsoft Windows 10.

2.3. Порівняння стійкості ФГ щодо класичних атак

Під час дослідження стійкості функцій гешування були використані вже відомі широкі результати досліджень [2, 4–6]. Ними підтверджено криптографічну стійкість вибраних ФГ до усіх відомих класичних атак.

2.4 Результати досліджень

В ході аналізу були отримані такі результати.

1. Усі ФГ успішно пройшли статистичні тести [10] (рис.5) з такими значеннями випадковості для NIST STS 800-22(2009 р.).

Отримано наступні показники складності (швидкодії) гешування (рис.6).

Геш-функція	Кількість тестів, що успішно пройдені на рівні α =0,99	Кількість тестів, що успішно пройдені на рівні α =0,96015
ГОСТ 34.11	126	188
SHA-2 256	130	188
SHA-2 512	135	187
SHA-3 256	133	187
SHA-3 512	126	186
Whirlpool	132	187
Kupyna 256	139	187
Kupyna 512	136	187

Рис. 5. Підсумкові середні дані статистичного тестування властивостей ФГ

На рис. 6 по осі Y відображена швидкість роботи у Мегабайтах/секунду. По осі X розташовані ФГ.

Отримані результати показують, що найкращі показники були отримані у ФГ SHA-2 та Kupyna. Найгірші результати швидкості серед цих алгоритмів виявились у SHA-3. Проте навіть ці дані можна покращити у подальших дослідженнях, адже реалізація була мовою програмування Java, яка дещо програє у швидкості мові C та C++ або мові програмування Assembler.

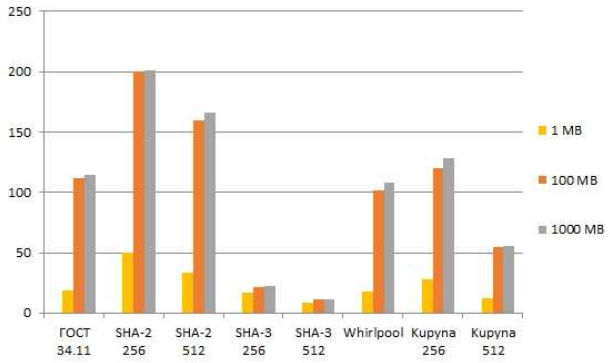

Рис. 6. Результати вимірів швидкості роботи швидкості ФГ

Для технології блокчейн швидкостей гешування за допомогою ФГ SHA-2 та Kupyna буде достатньо, адже середній розмір блоків у блокчейн мережах лише інколи перевищує 1МБ.

2. Результати аналізу стійкості ФГ наведені на рисунку 7.

Алгоритм	Безпека у бітах		
	Пошук колізії	Пошук прообразу	Пошук другого прообразу
ГОСТ 34.11-45	128 (256)	248 (512)	248 (512)
SHA-2 256	128	248	248
SHA-2 512	256	494	494
SHA-3 256	128	256	256
SHA-3 512	256	512	512
WHIRLPOOL	120	512	512
Kupyna-256	128	256	256
Kupyna-512	256	512	512

Рис. 7. Порівняльний аналіз класичних атак на ФГ

У цілому за результатами аналізу можна зробити висновок, що не дивлячись на зростання потужностей класичних комп'ютерів, сучасні функції гешування дозволяють забезпечити необхідний рівень стійкості проти усіх відомих атак.

Додатково була проаналізована загроза щодо створення та застосування квантового комп'ютера для здійснення атак на ФГ. Проведений аналіз показав, що квантовий комп'ютер здатний створити загрозу сучасним алгоритмам ФГ тільки з обмеженими розмірами параметрів, проте не на теперішньому етапі їх розвитку.

Було зроблено певний прогноз [11] щодо збільшення обчислювальної потужності квантового комп'ютера на найближчі 10 років із урахуванням необхідної для атаки потужності.

Припустимо, що кількість кубітів зростатиме експоненційно [12], тобто подвоюватиметься кожні 10 місяців, тоді як менш оптимістичне припущення передбачає подвоєння кожні 20 місяців. Ці дві екстраполяції наведені на рис. 8.

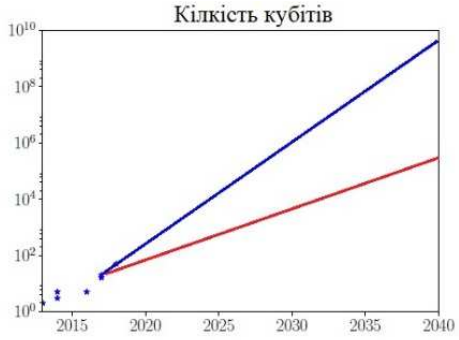

Рис. 8. Перспективи розвитку квантового комп'ютера

Далі, атака за допомогою методу Гровера [2] дозволяє зменшити час, необхідний для пошуку, наприклад, колізії функції гешування, до приблизно кореня із часу виконання класичної атаки:

$$O\left(\sqrt{\frac{N}{l}}\right).$$

Це дуже суттєве поліпшення, проте цього на даний момент не достатньо. На прикладі мережі Біткоїн, що використовує ФГ SHA-2 256 можна побачити, що потужності одного квантового комп'ютера не достатньо для суттєвого впливу на мережу.

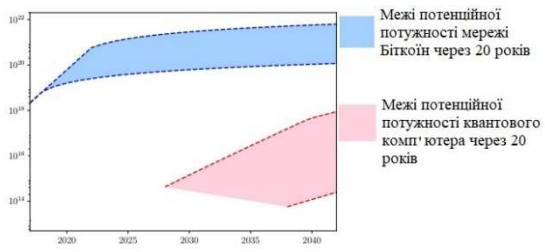

Рис. 9. Порівняння потужності мережі Біткоїн та квантового комп'ютера

Також була обчислена необхідна кількість кубітів для виконання атаки. Йде мова про фіксоване значення у 2402 логічних кубіти, що буде досягнуто згідно з прогнозами не раніше, ніж у 2027 році.

ВИСНОВКИ

На основі отриманих даних можна виділити таке:

1. Всі розглянуті стандартизовані ФГ пройшли статистичне тестування, тестування щодо складності гешування (швидкодії) та перевірку на криптографічну стійкість проти класичних атак.

3. Певну перевагу на сьогодні для використання при побудові блокчейн мереж є SHA-2, SHA-3 та Kupyna. Вказані і ФГ є стійкими проти класичного криптоаналізу, в тому числі: до знаходження прообразу; до знаходження другого прообразу та до виникнення чи створення колізій.

4. Більш конкретно можна відзначити SHA-2 та Kupyna, адже ці ФГ показують також найкращі результати щодо складності (швидкодії) гешування даних.

5. З огляду на потенціальну небезпеку квантового комп'ютера, доцільно використовувати ФГ із виходом у 512 біт: SHA-2 512, SHA-3 512 та Kupyna 512, адже це суттєво збільшує стійкість мережі, а втрати у швидкодії гешування є незначними.

6. Практично та швидще всього і теоретично, якщо розглядати атаку на основі методу Гровера ФГ при довжинах геш значень 512 бітів, а в перехідний період і при довжинах 256 бітів, забезпечуватимуть експоненційну складність навіть найбільш загрозливих атак на основі створення колізій.

7. Аналіз розвитку квантових комп'ютерів показав, що навіть за оптимістичного прогнозу, квантовий комп'ютер та відповідне математичне забезпечення можуть бути створені не раніше 2027 року.

8. На наш погляд, у ході оцінки можливостей та обґрунтування вибору ФГ для застосування в перспективних блокчейн мережам, можна рекомендувати до застосування ФГ із виходом у 512 біт: SHA-2 512, SHA-3 512 та Kupyna 512.

9. Безумовним є той факт, що в подальшому потрібно проводити дослідження властивостей та умов застосування вибраних ФГ, що рекомендуються до застосування.

Література

[1] *Andreas M.* Antonopoulos Mastering Bitcoin: Unlocking Digital Cryptocurrencies / Andreas M. Antonopoulos – K.: NGITS, 2014. – С. 10–150.

[2] За редакцією Горбенко І.Д. Горбенко Ю.І. «Побудування та аналіз систем, протоколів та засобів криптографічного захисту інформації». Монографія. Харків. Форт. 2015. – 902 с.

[3] *Don Tapscott, Alex Tapscott* Blockchain Revolution: How the Technology Behind Bitcoin is Changing Money, Business, and the World / Don Tapscott, Alex Tapscott Blockchain – K. : Information Systems, 2016 – С. 65–102.

[4] Криптографические хэш-функции [Електронний ресурс]. – Режим доступу: www/ URL: http://bit.nmu.org.ua/

ua/student/metod/cryptology/D0%BB%D0%B5%D0%BA%D1%86%D0%B8%D1%8F17.pdf – 04.10.2018 р.
[5] Возможные атаки на функции хэширования [Електронний ресурс]. – Режим доступу: www/ URL: https://studfiles.net/preview/2157418/page:2/ – 04.10.2018 р.
[6] Прикладна криптологія. Теорія. Практика. Застосування : монографія / І.Д. Горбенко, Ю.І. Горбенко; Міністерство освіти і науки, молоді та спорту України, ХНУРЕ, ПАТ "ІІТ" – Харків, 2012 – С. 340–347.
[7] Алгоритмы шифрования – основа работы криптовалют [Електронний ресурс]. – Режим доступу: www/ URL: https://tgraph.io/Algoritmy-shifrovaniya--osnova-raboty-kriptovalyut-09-27 – 14.09.2018 р.
[8] Comparison of cryptographic hash functions [Електронний ресурс]. – Режим доступу: www/ URL: https://en.wikipedia.org/wiki/Comparison_of_cryptographic_hash_functions – 17.10.2018 р.
[9] NISTIR 7896. Third-Round Report of the SHA-3 Cryptographic Hash Algorithm Competition / NIST, 2012. – С. 50–65.
[10] Analysis of the Kupyna-256 Hash Function / Christoph Dobraunig, Maria Eichlseder, and Florian Mendel, Graz University of Technology, – Austria. 2016.
[11] Квантовые компьютеры / Л. Федичкин, ФТИ РАН. НиЖ, 2001, – С. 20–33.
[12] Quantum search using Grover's algorithm [Електронний ресурс]. – Режим доступу: www/ URL: http://savepearlharbor.com/?p=222456 – 18.06.2014 р.

Надійшла до редколегії 26.11.2018

Кравчук Павло Вікторович, студент, ХНУРЕ. Галузь наукових інтересів – аналіз і тестування асиметричних перетворень, блокчейн технології.

Горбенко Іван Дмитрович, докт. техн. наук, професор кафедри безпеки інформаційних технологій (БІТ) Харківського національного університету радіоелектроніки (ХНУРЕ), академік Академії наук прикладної радіоелектроніки. Галузь наукових інтересів – створення, аналіз і реалізація систем і засобів захисту інформації; дослідження і реалізація криптографічних протоколів.

Пушкарьов Андрій Іванович, директор департаменту захисту інформації Адміністрації державної служби спеціального зв'язку та захисту інформації України. Галузь наукових інтересів – теорія захисту інформації, інформаційна та кібербезпека держави.

УДК 621.3.06

Кравчук П. В. **Анализ применения хеш-функций в технологии Blockchain** / П. В. Кравчук, И. Д. Горбенко, А. И. Пушкарев // Прикладная радиоэлектроника: науч. – техн. журнал. – 2018. – Том 17. № 3, 4. – С. 147–151.

В статье представлены результаты анализа выбранных хеш-функций для их применения в системах, использующих технологию Blockchain. Кроме того, был проведен сравнительный анализ их стойкости, скорости и противодействие различным атакам; сформированы правила применения хеш-функций.

Ключевые слова: хеш - значения, электронная подпись, криптографические механизмы, криптографическая стойкость, услуги безопасности, технологии блокчейн, сложность криптопреобразований, функция хеширования.

Ил.: 9. Библиогр.: 12 назв.

UDC 621.3.06

Kravchuk P. V. **Analysis of the application of Hash functions in Blockchain technology** / P. V. Kravchuk, I. D. Gorbenko, A. I. Pushkarev // Applied Radio Electronics: Sci. Journ. – 2018. – Vol. 17. № 3, 4. – P. 147–151.

The paper presents the results of the analysis of selected hash functions for their application in systems using Blockchain technology. In addition, the resuls of a comparative analysis of their stiffness, speed, their defense against various attacks were made; the rules of application of hash functions were formed.

Keywords: hash-value, electronic signature, cryptographic mechanisms, cryptographic stability, security services, blockchain technology, complexity of cryptotransformations, hashing function.

Fig.9. Ref.: 12 items.