ХЕШИРОВАНИЕ КАК ОДИН ИЗ СПОСОБОВ ЗАЩИТЫ ИНФОРМАЦИИ

Слепцов Д.В.

Федеральное государственное бюджетное образовательное учреждение высшего образования «Брянский государственный университет имени академика И. Г. Петровского»

г. Брянск, Россия

***Аннотация**: В статье рассматривается один из современных методов криптографических преобразований массивов данных – хеширование, незаменимый и повсеместно распространенный инструмент защиты информации.*

***Ключевые слова**: криптография, хеширование, хеш-функция, алгоритм хеширования.*

HASHING AS A WAY OF PROTECTING INFORMATION

Sleptsov D.V.

Federal State-Funded Educational Institution of Higher Education "Bryansk State University named after Academician I.G. Petrovsky"

Bryansk, Russia

***Annotation**: The article deals with one of the modern methods of cryptographic transformation of amounts of data - hashing, an indispensable and widespread tool for protecting information.*

***Keywords**: cryptography, hashing, hash function, hash algorithm.*

Сегодня нет необходимости обосновывать актуальность угроз целостности и конфиденциальности информации. Еще пару десятков лет назад задача обеспечения безопасности информации решалась при помощи использования межсетевых экранов и разграничения доступа. Сейчас этих технологий недостаточно, поскольку любая информация, которая имеет финансовую, посредническую, военную, политическую, научную ценность,

хранится в компьютере и, следовательно, подвергается угрозе. Дополнительным риском становится возможность перехвата управления критическими объектами информационной инфраструктуры. Согласно статистическим данным, более 80% компаний несут финансовые убытки из-за нарушения целостности и конфиденциальности используемых данных [1].

Наука о методах обеспечения конфиденциальности (невозможности прочтения информации посторонним), целостности данных (невозможности незаметного изменения информации), аутентификации (проверки подлинности авторства или иных свойств объекта), а также невозможности отказа от авторства называется криптография. Криптография – одна из старейших наук, её история насчитывает несколько тысяч лет [5].

Изначально криптография изучала методы шифрования информации – обратимого преобразования исходного текста на основе секретного алгоритма или ключа в шифрованный текст. Традиционная криптография образует раздел симметричных криптосистем, в которых шифрование и дешифрование проводится с использованием одного и того же секретного ключа. Помимо этого раздела современная криптография включает в себя асимметричные криптосистемы, системы электронной цифровой подписи (ЭЦП), хеш-функции, управление ключами, получение скрытой информации, квантовую криптографию.

Одним из интересных и перспективных направлений современной криптографии является раздел, изучающий хеш-функции. Хеш-функцией называется алгоритм, конвертирующий строку произвольной длины (сообщение) в битовую строку фиксированной длины, называемой хеш-кодом, проверочной суммой или цифровым отпечатком [2].

Во многих технологиях безопасности применяются односторонние функции шифрования, называемые также хеш-функциями. Основное назначение подобных функций – получение из сообщения произвольного размера его хеш-кода – значения фиксированного размера. Хеш-код может быть использован в качестве контрольной суммы исходного сообщения,

обеспечивая таким образом (при использовании соответствующего протокола) контроль целостности информации.

Основными свойствами хеш-функции являются:

-на вход хеш-функции подается сообщение произвольной длины;

-на выходе хеш-функции формируется блок данных фиксированной длины;

-значения на выходе хеш-функции распределены по равномерному закону;

-при изменении одного бита на входе хеш-функции существенно изменяется выход [4].

Последнее свойство хеш-функций позволяет применять их в следующих случаях:

-при построении ассоциативных массивов;

-при поиске дубликатов в сериях наборов данных;

при построении уникальных идентификаторов для наборов данных;

-при вычислении контрольных сумм от данных для последующего обнаружения в них ошибок, возникающих при хранении и/или передаче данных;

-при сохранении паролей в системах защиты в виде хеш-кода (для восстановления пароля по хеш-коду требуется функция, являющаяся обратной по отношению к использованной хеш-функции);

-при выработке электронной подписи [3].

В общем случае нет однозначного соответствия между исходными (входными) данными и хеш-кодом (выходными данными). Возвращаемые хеш-функцией значения (выходные данные) менее разнообразны, чем значения входного массива (входные данные). Случай, при котором хеш-функция преобразует несколько разных сообщений в одинаковые хеш-коды, называется «коллизией». Вероятность возникновения коллизий используется для оценки качества хеш-функций.

Алгоритмы хеширования характеризуются разрядностью, вычислительной сложностью, криптостойкостью.

Хорошая хеш-функция должна удовлетворять двум свойствам:

-быстро вычисляться;

-минимизировать количество коллизий.

Предположим, для определённости, что K - количество ключей, а хеш-функция h(k) имеет не более M различных значений: для любого k, выполняется неравенство $0 \le h(k) < M$.

Существует несколько простых и надежных методов, на которых базируются многие хеш-функции [4].

Первый метод заключается в том, что мы используем в качестве хеша остаток от деления на M, где M - это количество всех возможных хешей:

$h(K) = K \bmod M$

При этом очевидно, что при чётном M значение функции будет чётным, при чётном K, и нечётным – при нечётном, что может привести к значительному смещению данных в файлах. На практике обычно выбирают простое.

Ещё следует сказать о методе хеширования, основанном на делении на полином по модулю два. В данном методе M также должна являться степенью двойки, а бинарные ключи K представляются в виде полиномов. В этом случае в качестве хеш-кода берутся значения коэффициентов полинома, полученного как остаток от деления K на заранее выбранный полином P степени m:

При правильном выборе P такой способ гарантирует отсутствие коллизий между почти одинаковыми ключами.

Второй метод состоит в выборе некоторой целой константы A, взаимно простой с w, где w - количество представимых машинным словом значений. Тогда можем взять хеш-функцию вида:

$$h(K) = \left[M \left\lfloor \frac{A}{w} * K \right\rfloor \right].$$

В этом случае, на компьютере с двоичной системой счисления, M является степенью двойки и h(K) будет состоять из старших битов правой половины произведения A*K.

Одной из вариаций данного метода является хеширование Фибоначчи, основанное на свойствах золотого сечения.

Хеширование Пирсона - алгоритм, предложенный Питером Пирсоном для процессоров с 8-битными регистрами, задачей которого является быстрое вычисление хеш-кода для строки произвольной длины. На вход функция получает слово W, состоящее из n символов, каждый размером 1 байт, и возвращает значение в диапазоне от 0 до 255. При этом значение хеш-кода зависит от каждого символа входного слова.

Среди преимуществ алгоритма следует отметить:

-простоту вычисления;

-не существует таких входных данных, для которых вероятность коллизии наибольшая;

-возможность модификации в идеальную хеш-функцию.

Идеальной хеш-функцией называется такая функция, которая отображает каждый ключ из набораS в множество целых чисел без коллизий. В математических терминах это инъективное отображение.

Функция $h(k):U \to [m]$ называется идеальной хеш-функцией для $S \subseteq U$, если она инъективна на S.

Функция $h(k):U \to [m]$ называется минимальной идеальной хеш-функцией для $S \subseteq U$, если она является ИХФ и $m = n = |S|$.

Для целого $k \geq 1$, функция $h(k):U \to [m]$ называется k-идеальной хеш-функцией для $S \subseteq U$ если для каждого$j \in [m]$ имеем $|\{x \in S | h(x) = j\}| \leq k$.

Идеальное хеширование применяется в тех случаях, когда необходимо присвоить уникальный идентификатор ключу, без сохранения какой-либо информации о ключе. Например, идеальное хеширование используется для ускорения работы алгоритмов на графах, в тех случаях, когда представление графа не умещается в основной памяти.

Универсальным хешированием называется хеширование, при котором используется не одна конкретная хеш-функция, а происходит выбор из заданного семейства по случайному алгоритму. Использование универсального хеширования обычно обеспечивает низкое число коллизий. Универсальное хеширование имеет множество применений, например, в реализации хеш-таблиц и криптографии.

Таким образом, хеширование – незаменимый и повсеместно распространенный инструмент, используемый для выполнения целого ряда задач, включая аутентификацию, проверку целостности данных, защиту файлов и даже обнаружение зловредного ПО. Существует масса алгоритмов хеширования, отличающихся криптостойкостью, сложностью, разрядностью и другими свойствами, которые сегодня используется во многих областях информационных технологий.

Список литературы

1. Алексенко О.Ю. Актуальность проблемы защиты информации [Электронный ресурс]. – Режим доступа: https://nsportal.ru/npo-spo/elektronnaya-tekhnika-radiotekhnika-i-svyaz/library/2017/11/22/aktualnost-problemy-zashchity

2. Брассар Ж. Современная криптология / Ж. Брассар – М.: ПОЛИМЕД, 1999. – 178 с.

3. Васильева И.Н. Криптографические методы защиты информации / И.Н. Васиева – М.: Юрайт, 2016. – 340 с.

4. Шнайер Б. Прикладная криптография/ Б. Шнайер – М.: ТРИУМФ, 2003. – 816 с.

5. Ященко В.В. Введение в криптографию / В.В. Ященко – М.: Издательство МЦНМО, 2012. – 342 с.