Кафедра информационной безопасности

Раздел №3:

«Криптография в сетях связи»

Тема №1:

«Аутентификация и цифровая подпись»

Занятие №2

«Криптографические хеш-функции»

Учебные и воспитательные цели:

1. Рассмотреть применение хэш-функций в обеспечении целостности данных.

2. Усвоить особенности применения ключевых и бесключевых функций хэширования.

3. Изучить стандарты хэш-функций.

4. Активизировать на изучение вопросов обеспечения безопасности информации криптографическими методами.

Учебные вопросы:

1. Функции хеширования и целостность данных

2. Ключевые и бесключевые функции хеширования

3. Стандарты хеш-функций

Первый учебный вопрос:

«Функции хеширования и целостность данных»

КОНТРОЛЬ УСВОЕНИЯ МАТЕРИАЛА

Контрольные вопросы для проверки усвояемости материала предыдущего занятия:

1. На какие группы могут быть разбиты алгоритмы идентификации?
2. В чем состоят недостатки систем с фиксированными паролями?
3. За счет чего повышается надежность идентификации при использовании пластиковой карты и личного идентификационного номера?
4. Каковы возможные схемы использования одноразовых паролей?
5. Для каких целей используется временная метка в протоколе типа "запрос-ответ"?
6. Чем могут быть заменены временные метки в протоколах типа "запрос-ответ"?
7. Какая идея лежит в основе протоколов с нулевым разглашением?
8. Какие типы атак могут быть использованы при нападении на протоколы идентификации?

ПОНЯТИЕ ХЕШ-ФУНКЦИЙ

Опр. 1. **Хеш-функция – функция, отображающая входное слово конечной длины в конечном алфавите в слово заданной, обычно фиксированной длины.**

Хеш-функции — это функции, предназначенные для "сжатия" произвольного сообщения или набора данных, записанного, как правило, в двоичном алфавите, в некоторую битовую комбинацию фиксированной длины, называемую *сверткой.*

В криптографии хеш-функции применяются для решения следующих задач:

1. построения систем контроля целостности данных при их передаче или хранении,

2. аутентификации источника данных.

[1] При решении первой задачи для каждого набора данных вычисляется значение хеш-функции (называемое *кодом аутентификации сообщения* или *имитовставкой*), которое передается или хранится вместе с самими данными. При получении данных пользователь вычисляет значение свертки и сравнивает его с имеющимся контрольным значением. Несовпадение говорит о том, что данные были изменены.

[!] Хеш-функция, служащая для выработки имитовставки, должна позволять (в отличие от обычной контрольной суммы) осуществлять обнаружение не только случайных ошибок в наборах данных, возникающих при их хранении и передаче, но и сигнализировать об активных атаках злоумышленника, пытающегося осуществить навязывание ложной информации.

[1] ***построения систем контроля целостности данных при их передаче или хранении*** **(продолжение)**

Опр. 2. **Ключевая хеш-функция – криптографическая хеш-функция, реализуемая кодом аутентификации, и предназначенная для обеспечения невозможности для противника и/или нарушителя создать новые или модифицировать передаваемые (или хранимые) сообщения.**

Имитовставки, формируемые с помощью ключевых хеш-функций, не должны позволять противнику создавать сфабрикованные сообщения *(fabrication)* при атаках типа *"имитация"* *(impersonation)* и модифицировать передаваемые сообщения *(modification)* при атаках типа *"подмена"* *(substitution).*

[2] **При решении второй задачи** — аутентификации источника данных — мы имеем дело с не доверяющими друг другу сторонами. В такой ситуации применяют схемы цифровой подписи, позволяющие осуществлять аутентификацию источника данных. Как правило, при этом сообщение, прежде чем быть подписано личной подписью, основанной на секретном ключе пользователя, "сжимается" с помощью хеш-функции, выполняющей функцию кода обнаружения ошибок.

Основными требованиями к ней являются гарантии невозможности подмены подписанного документа, а также подбора двух различных сообщений с одинаковым значением хеш-функции.

Обозначим через X множество, элементы которого будем называть сообщениями. Пусть Y — множество двоичных векторов фиксированной длины.

Опр. 3. **Хеш-функцией называется всякая функция $h: X \to Y$ легко вычислимая и такая, что для любого сообщения M значение $h(M) = H$ *(свертка)* имеет фиксированную битовую длину.**

[2] *аутентификации источника данных* (продолжение)

Как правило, хеш-функции строят на основе так называемых *одношаговых сжимающих функций* $y = f(x_1, x_2)$ двух переменных, где x и y — двоичные векторы длины m и n соответственно, причем n — длина свертки. Для получения значения $h(M)$ сообщение M сначала разбивается на блоки длины m, а затем к полученным блокам $M_1, M_2, .., M_N$ применяют следующую последовательную процедуру вычисления свертки:

$$H_0 = v,$$
$$H_i = f(M_i, H_{i-1}),\ i = 1, ..., N, \qquad (1)$$
$$h(M) = H_N.$$

Особо выделяют два важных типа криптографических хеш-функций — *ключевые* [1] и *бесключевые* [2].

[1] Ключевые применяются в системах с симметричными ключами. Ключевые хеш-функции называют ***кодами аутентификации сообщений* (КАС)** (*message authentication code* **(MAC)**). Они дают возможность без дополнительных средств гарантировать как правильность источника данных, так и целостность данных в системах с доверяющими друг другу пользователями.

[2] Бесключевые хеш-функции называются ***кодами обнаружения ошибок*** (*modification detection code* **(MDC)** *или manipulation detection code, message integrity code* **(MIC)**). Они дают возможность с помощью дополнительных средств (например, шифрования, использования защищенного канала или цифровой подписи) гарантировать целостность данных. Эти хеш-функции могут применяться в системах как с доверяющими, так и не доверяющими друг другу пользователями.

Второй учебный вопрос:

«Ключевые и бесключевые функции хеширования»

2.1. Ключевые хеш-функции

В криптографических приложениях к ключевым функциям хеширования предъявляются следующие основные требования:

- *невозможность фабрикации* [1]

(означает высокую сложность подбора сообщения с правильным значением свертки);

- *невозможность модификации* [2]

(означает высокую сложность подбора для заданного сообщения с известным значением свертки другого сообщения с правильным значением свертки).

Ключевые функции применяются в ситуациях, когда стороны доверяют друг другу и могут иметь общий секретный ключ.

В качестве примера рассмотрим широко распространенную хеш-функцию, построенную на основе одношаговой сжимающей функции вида

$$f_k(x, H) = E_k(x \oplus H)$$

где E_k — алгоритм блочного шифрования.

Алгоритм вычисления свертки имеет следующий вид:

$$H_0 = 0,$$
$$H_i = E_k(M_i \oplus H_{i-1}), \quad i = 1, \ldots, N, \quad (2)$$
$$h(M) = H_N.$$

Данный алгоритм фактически совпадает с режимом шифрования со сцеплением блоков, с той лишь разницей, что в качестве результата берется не весь шифртекст $H_1, H_2, \ldots, H_N$, а только его последний блок. Такой режим в ГОСТе 28147-89 называется ***режимом выработки имитовставки***.

2.2. Бесключевые хеш-функции

Обычно требуется, чтобы бесключевые хеш-функции обладали следующими свойствами:

- *Однонаправленность* [1]
  (означает высокую сложность нахождения сообщения с заданным значением свертки);
- *устойчивость к коллизиям* [2]
  (означает высокую сложность нахождения пары сообщений с одинаковыми значениями свертки);
- *устойчивость к нахождению второго прообраза* [3]
  (означает высокую сложность нахождения второго сообщения с тем же значением свертки для заданного сообщения с известным значением свертки).

Хеш-функция контрольной суммы CRC-32 является линейным отображением и поэтому не удовлетворяет ни одному из этих трех свойств. Использование в качестве бесключевой хеш-функции рассмотренной выше ключевой **(2)**, также нецелесообразно, так как обратимость блочного шифрования позволяет подбирать входное сообщение для любого значения свертки при фиксированном и общеизвестном ключе.

Для построения примера хеш-функции, удовлетворяющей свойству [1], рассмотрим функцию, заданную формулой $\mathbf{g_k(x) = E_k(x) \oplus x}$, где E_k — алгоритм блочного шифрования. Такая функция является однонаправленной по обоим аргументам. Поэтому на ее основе можно построить хеш-функцию по правилу [1], определив одношаговую сжимающую функцию одной из следующих формул:

$$\mathbf{f(x, H) = E_H(x) \oplus x}$$

или

$$\mathbf{f(x, H) = E_x(H) \oplus H}$$

Первая из этих функций лежит в основе ***российского стандарта хеш-функции (ГОСТ Р 34.11-94)***, а вторая — в основе ***американского стандарта SHA***.

Третий учебный вопрос:

«Стандарты хеш-функций»

Однонаправленные хеш-функции на основе симметричных блочных алгоритмов

Если принять, что получаемая хеш-функция корректна, безопасность схемы хеширования базируется на безопасности лежащего в ее основе блочного алгоритма.

Схема хеширования, у которой длина хеш-значения равна длине блока, показана на рис. 1. Ее работа описывается выражениями:

$$H_0 = I_H,$$
$$H_i = E_A(B) \oplus C,$$

где I_H – некоторое случайное начальное значение; **A**, **B** и **C** могут принимать значения M_i, H_{i-1}, $(M_i \oplus H_{i-1})$ или быть константами.

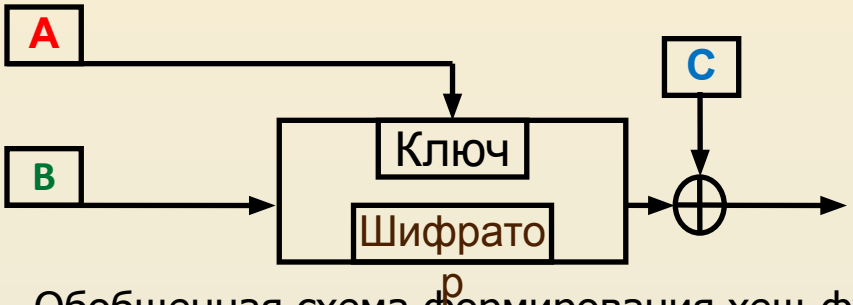
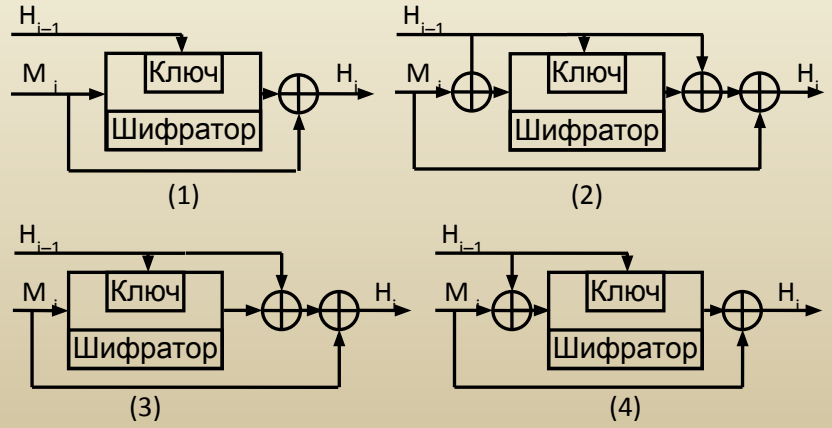


Рис. 1. Обобщенная схема формирования хеш-функции

Первые четыре схемы хеширования, являющиеся безопасными при всех атаках, приведены на рис. 2.



Отечественный стандарт хеш-функции

Российский стандарт ГОСТ Р 34.11-94 определяет алгоритм и процедуру вычисления хеш-функции для любых последовательностей двоичных символов, применяемых в криптографических методах обработки и защиты информации. Этот стандарт базируется на блочном алгоритме шифрования ГОСТ 28147-89 с 64-битовым блоком и 256-битовым ключом и формирует 256-битовое хеш-значение.

Функция сжатия $H_i = f(M_i, H_{i-1})$ (оба операнда M_i и H_{i-1} являются 256-битовыми величинами) определяется следующим образом:

1. Генерируются 4 ключа шифрования K_j, $j = 1…4$, путем линейного смешивания M_i, H_{i-1} и некоторых констант C_j.
2. Каждый ключ K_j, используют для шифрования 64-битовых подслов h_i слова H_{i-1} в режиме простой замены: $S_j = E_{kj}(h_j)$. Результирующая последовательность S_4, S_3, S_2, S_1 длиной 256 бит запоминается во временной переменной S.
3. Значение H_i является сложной, хотя и линейной функцией смешивания S, M_i и H_{i-1}.

При вычислении окончательного хеш-значения сообщения M учитываются значения трех связанных между собой переменных:

H_n – хеш-значение последнего блока сообщения;

Z – значение контрольной суммы, получаемой при сложении по модулю 2 всех блоков сообщения;

L – длина сообщения.

Эти три переменные и дополненный последний блок M_n сообщения объединяются в окончательное хеш-значение следующим образом:

$$H = f(Z \oplus M_n, f(L, f(M_n, H_n))).$$

Данная хеш-функция определена стандартом ГОСТ Р 34.11-94 для использования совместно с российским стандартом электронной цифровой подписи.