О криптографических свойствах алгоритмов, сопутствующих применению стандартов ГОСТ Р 34.10-2012 и ГОСТ Р 34.11-2012

Смышляев С.В.
Алексеев Е.К., Ошкин И.Б., Попов В.О.

Рассмотрим следующие сопутствующие применению стандартов ГОСТ Р 34.10-2012 и ГОСТ Р 34.11-2012 алгоритмы.

- $\mathrm{HMAC_t}$, $\mathrm{t} \in \{256, 512\}$, — коды аутентификации сообщений на основе хэш-функций;
- $\mathrm{PRFTLS_t}$, $\mathrm{t} \in \{256, 512\}$, — псевдослучайные функции, аналогичные используемым в протоколе TLS;
- $\mathrm{PRFKEYMAT_t}$, $\mathrm{t} \in \{256, 512\}$, — псевдослучайные функции, аналогичные используемым в протоколах IPsec для порождения ключевого материала;
- $\mathrm{PRFPLUS_t}$, $\mathrm{t} \in \{256, 512\}$, — псевдослучайные функции, аналогичные используемым в протоколе IKEv2;

- KDFTREE — семейство функций диверсификации;
- KDF — функция диверсификации, предназначенная для однократной выработки ключевого материала по исходному ключу;
- $\mathrm{VKO_t}$, $\mathrm{t} \in \{256, 512\}$, — функции ключевого обмена на основе протокола Диффи–Хеллмана;

В настоящей работе проводится анализ защищенности данных сопутствующих алгоритмов в предположении положительных криптографических свойств алгоритма хэширования ГОСТ Р 34.11-2012, а также в предположении вычислительной сложности задач Диффи–Хеллмана (распознавательной и вычислительной) в рекомендованных для использования группах точек эллиптических кривых.

Соглашения и обозначения

- $\mathbb{N} = \{1, 2, \dots\}$ — натуральные числа, $\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$;
- $\mathbb{F}_2$ — поле $\mathrm{GF}(2)$;
- V_n — множество наборов длины $n \in \mathbb{N}_0$ с элементами из $\mathbb{F}_2$ (векторное пространство над $\mathbb{F}_2$ размерности n);
- $\mathrm{V}^* = \bigcup_{r \in \mathbb{N}_0} \mathrm{V}_r$ — множество битовых строк, имеющих произвольные конечные длины;
- Для $t \in \{256, 512\}$ через GH_t будем обозначать хэш-функцию ГОСТ Р 34.11-2012 с длиной выхода равной t битам ($\mathrm{GH}_t : \mathrm{V}^* \to \mathrm{V}_t$);
- $[m]_2$ — байтовое представление произвольного числа $m \in \mathbb{N}$.

Определения

Под противником, которого будем обозначать через Adv, далее понимается вероятностная машина Тьюринга.

Определение

Некоторый объект (функцию или задачу) будем называть (ε, T)-стойким в некоторой модели противника (если речь идет о задаче, то модель может быть не указана), если в условиях указанной в модели атаки противник не может реализовать соответствующую угрозу с вероятностью успеха (или преобладанием для задач распознавания) большей или равной ε, затратив менее T операций.

КРИПТО ПРО

Определение

Распознавательной задачей Диффи–Хеллмана (DDH) (см. D. Boneh, The Decision Diffie–Hellman Problem) называется следующая задача: на вход подаются точки xP, yP и Q, выбранные случайно и равновероятно из группы E точек эллиптической кривой; требуется выдать 1, если $Q = xyP$, и 0 иначе.

Распознавательная задача Диффи–Хеллмана считается труднорешаемой для групп точек эллиптических кривых простого порядка (см., например, R.Gennaro, H.Krawczyk and T.Rabin, Secure Hashed Diffie-Hellman over Non-DDH Groups).

КРИПТО ПРО

HMAC

Длина n ключа K для функций HMAC_t ($t \in \{256, 512\}$) должна удовлетворять условию $256 \leqslant n \leqslant 512$. Для алгоритма HMAC_{512} рекомендуется использовать ключи размером 512.
Процедура дополнения ключа (общая для HMAC_{256} и HMAC_{512}):

$$K^* = K|0^{512-n} \in V_{512}.$$

$\mathrm{HMAC}_{256} : V_{[256,512]} \times V^* \to V_{256}$. Если $T \in V_m$ и $K \in V_n$, то

$$\mathrm{HMAC}_{256}(K, T) = \mathrm{GH}_{256}(K^* \oplus \mathrm{opad}|\mathrm{GH}_{256}(K^* \oplus \mathrm{ipad}|T)).$$

$\mathrm{HMAC}_{512} : V_{[256,512]} \times V^* \to V_{512}$. Если $T \in V_m$ и $K \in V_n$, то

$$\mathrm{HMAC}_{512}(K, T) = \mathrm{GH}_{512}(K^* \oplus \mathrm{opad}|\mathrm{GH}_{512}(K^* \oplus \mathrm{ipad}|T)),$$

где $\mathrm{opad}, \mathrm{ipad}$ — фиксированные строки из V_{512}.

КРИПТО ПРО

PRFTLS

Пусть $A(secret, seed) =$

$$\underbrace{seed}_{A_0} | \underbrace{HMAC_t(secret, A(0))}_{A(1)} | \underbrace{HMAC_t(secret, A(1))}_{A(2)} | \dots .$$

$PRFTLS_t(secret, label, seed) = T_1|T_2|T_3|\dots$, где

$T_i = HMAC_t(secret, A(i)|label|seed)$ для $i = 1, 2, \dots .$

Функции $PRFTLS_t$ определены по аналогии с используемыми в протоколе TLS.

PRFKEYMAT

$PRFKEYMAT_t(K, S) =$

$$\underbrace{HMAC_t(K, S)}_{T(1)} | \underbrace{HMAC_t(K, T(1)|S)}_{T(2)} | \underbrace{HMAC_t(K, T(2)|S)}_{T(3)} | \dots .$$

Функции $PRFKEYMAT_t$ по схеме задания аргументов в итерациях аналогичны функции KEYMAT в протоколе IKE.

КРИПТО ПРО

PRFPLUS

$$\mathrm{PRFPLUS_t(K,S)} = \underbrace{\mathrm{HMAC_t(K,S|01)}}_{\mathrm{T(1)}}|\underbrace{\mathrm{HMAC_t(K,T(1)|S|02)}}_{\mathrm{T(2)}}|\underbrace{\mathrm{HMAC_t(K,T(2)|S|03)}}_{\mathrm{T(3)}}|\ldots.$$

Функции $\mathrm{PRFPLUS_t}$ по схеме задания аргументов в итерациях аналогичны функции prf+ в IKE v.2.

KDFTREE

$\mathrm{KDFTREE(K,label,seed,R)} = \mathrm{K(1)|K(2)|K(3)|\ldots}$, где $\mathrm{K(i)} = \mathrm{HMAC_{256}(K,[i]_2|label|00|seed|[L]_2)}$, R — внешний фиксируемый параметр, $\mathrm{R} \in \{1,2,3,4\}$, L — необходимая битовая длина вырабатываемого ключевого материала.

Функция диверсификации KDFTREE использует общие принципы задания входных параметров и выхода для функций диверсификации, изложенные в документе NIST SP 800-108.

KDF

$\mathrm{KDF}(\mathrm{K}, \mathrm{label}, \mathrm{seed}) = \mathrm{KDFTREE}(\mathrm{K}, \mathrm{label}, \mathrm{seed}, 256) = \mathrm{HMAC}_{256}(\mathrm{K}, 01|\mathrm{label}|00|\mathrm{seed}|01|00)$.

Заметим, что функция KDF является частным случаем функции KDFTREE: длина выхода фиксирована и равна 256 битам (один блок).

VKO

$\mathrm{VKO}_t(\mathrm{x}, \mathrm{y}, \mathrm{UKM}) = \mathrm{GH}_t((\mathrm{UKM} \cdot \mathrm{x})(\mathrm{mod}\ \mathrm{q}) \cdot (\mathrm{yP}))$.

Алгоритм VKO_t могут использоваться для согласования ключей ГОСТ Р 34.10-2012, t бит.

Алгоритм определен по аналогии с разделом 5.2 RFC 4357.

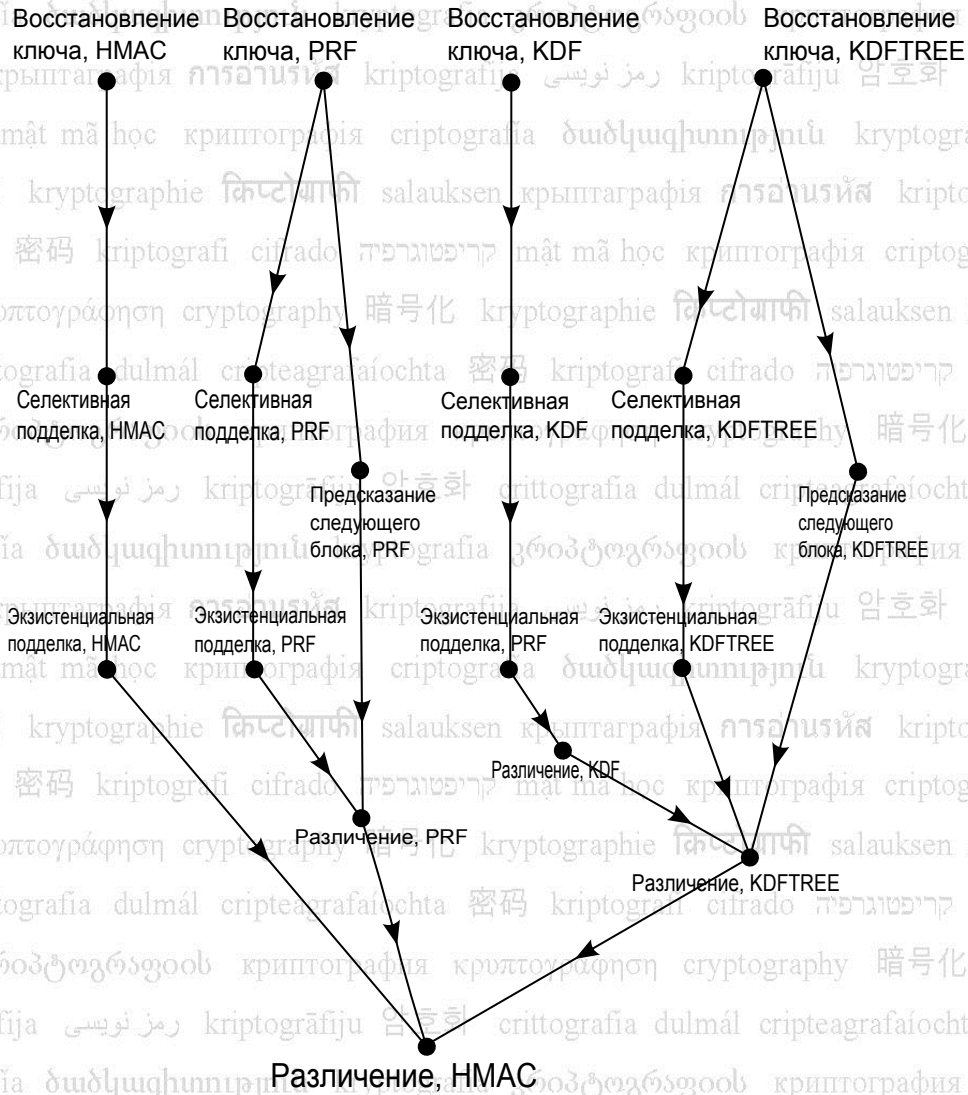
Обоснование стойкости алгоритма HMAC

Для функции $F(K, M, r) : V_n \times V_m \times \mathbb{N} \to V_{r \cdot t}$ ($F(K, M, r) = T_1|T_2|\ldots|T_r$) рассматриваются следующие угрозы:

- Восстановление ключа: требуется найти ключ K;
- Селективная подделка: требуется построить $D = F(K, M, r)$ (значение r фиксировано) для некоторого фиксированного M, выбранного противником перед началом работы;
- Экзистенциальная подделка: требуется построить такие M и D, что $D = F(K, M, r)$ (значение r фиксировано);
- Подделка следующего блока: для произвольных фиксированных M и s найти T_{s+1} по $T_1, \ldots, T_s$.
- Различение: для неизвестного фиксированного ключа K отличить функцию F (параметр r фиксирован) от функции, случайно равновероятно выбранной из множества всех функций из V_m в $V_{r \cdot t}$.

Для функции $F(K, M, r)$ рассматривается следующая атака:

- Атака с адаптивным выбором открытых текстов: противник может выбирать сообщения M_i, зная все выбранные им ранее сообщения и результаты работы оракула на этих сообщениях.



КРИПТО ПРО

Определение

Функция $F(K, m)$ называется (ε, T, q, n)-стойкой к коллизиям в слабом смысле, если для любого (T, q, n)-противника, производящего запросы m длины не более n блоков к оракулу O_{wcr}, который вычисляет функцию $F(K, \cdot)$, вероятность найти коллизию ($m \neq m'$ такие, что $F(K, m) = F(K, m')$) меньше или равна ε.

Определение

Функция $F(K, m)$ называется (ε, T, q, n)-стойкой псевдослучайной функцией, если для любого (T, q, n)-противника, производящего запросы m длины не более n блоков к оракулу O_F, который возвращает значения либо случайной функции, либо $F(K, \cdot)$, преобладание меньше или равно ε.

GH′: ограничение функции GH на входы, состоящие из двух блоков.

Определение

Пару функций (GH, GH') будем называть (ε, T, q, n)-стойкой против атаки с (r', r'')-связанными ключами, где $r'(K) = K \oplus ipad$, а $r''(K) = K \oplus opad$, если для любого (T, q, n)-противника, использующего оракул O_{rka}, преобладание меньше или равно ε.

O_{rka} выбирает случайное значение $l \in \{0, 1\}$.

- Если $l = 0$, случайно и равновероятно выбирает ключ K. На запрос $(m, 1)$, где $m \in V^*$, выдает $GH(r'(K)|m)$, а на запрос $(m, 2)$, где $m \in V_{512}$, — $GH'(r''(K)|m)$.
- Если $l = 1$, случайно, равновероятно и независимо выбирает ключи K_1 и K_2. На запрос $(m, 1)$ отвечает $GH(K_1|m)$, а на запрос $(m, 2)$ — $GH'(K_2|m)$.

КРИПТО ПРО

Теорема

Если функция GH является (ε_1, T, q, n)-стойкой к коллизиям, функция GH′ является $(\varepsilon_2, T + nqT_g, q, 1)$-стойкой псевдослучайной функцией, а (GH, GH') является $(\varepsilon_3, T, 2q, n)$-стойкой против атаки с (r', r'')-связанными ключами, то HMAC является $(\varepsilon_1 + \varepsilon_2 + \varepsilon_3, T, q, n)$-стойкой псевдослучайной функцией.

Замечание

Использование в формулировке теоремы пары функций (GH, GH') объясняется тем, что именно они участвуют в вычислении целевой для обоснования функции HMAC. Использование чуть более естественной пары (GH, GH) привело бы к неоправданному усилению противника.

Обоснование стойкости алгоритма VKO

VKO

$\mathrm{VKO_t(x, y, UKM)} = \mathrm{GH_t((UKM \cdot x)(mod\ q) \cdot (yP))}$.

Алгоритм $\mathrm{VKO_t}$ могут использоваться для согласования ключей ГОСТ Р 34.10-2012, t бит.

Алгоритм определен по аналогии с разделом 5.2 RFC 4357.

Модель 1.

Атака: противнику известны UKM, xP, yP, K = VKO(x, y, UKM) и некоторая информация $R_1(x, y)$ о паре (x, y).
Угроза: противник узнает некоторую новую информацию $R_2(x, y)$ о паре (x, y).

Заметим, что Модель 1 на самом деле является семейством моделей, которое параметризовано парой функций R_1, R_2 и распределением $\mathcal{D}$. Поясним смысл значений $R_1(x, y)$ и $R_2(x, y)$ на примерах.

Пример

Для пассивного противника, имеющего доступ лишь к каналу связи, $R_1(x, y) = 0$, а $R_2(x, y) = x$.

Пример

Пусть противник получает информацию о x по побочному каналу. В этом случае $R_1(x, y)$ — информация об x, полученная по побочному каналу.

Модель 1.

Атака: противнику известны UKM, xP, yP, K = VKO(x, y, UKM) и некоторая информация $R_1(x, y)$ о паре (x, y).
Угроза: противник узнает некоторую новую информацию $R_2(x, y)$ о паре (x, y).

Заметим, что Модель 1 на самом деле является семейством моделей, которое параметризовано парой функций R_1, R_2 и распределением $\mathcal{D}$. Поясним смысл значений $R_1(x, y)$ и $R_2(x, y)$ на примерах.

Пример

Для пассивного противника, имеющего доступ лишь к каналу связи, $R_1(x, y) = 0$, а $R_2(x, y) = x$.

Пример

Пусть противник получает информацию о x по побочному каналу. В этом случае $R_1(x, y)$ — информация об x, полученная по побочному каналу.

КРИПТО ПРО

Обоснование (идейно)

- Предположим, что существует такой алгоритм A, который для некоторых произвольных фиксированных отображений $R_1(x, y), R_2(x, y)$ может по $R_1(x, y)$, xP, yP, UKM и $VKO(x, y, UKM)$ эффективно вычислять $R_2(x, y)$ с вероятностью p, которая не является пренебрежимо малой.
  - Поскольку хэш-функция H является эффективно вычислимой, то с помощью этого алгоритма можно построить эффективный алгоритм, который вычисляет $R_2(x, y)$ с не меньшей вероятностью успеха по $R_1(x, y)$, xP, yP и xyP.
  - Поэтому стойкость алгоритма VKO в первой модели противника не меньше стойкости алгоритма Диффи–Хеллмана.

КРИПТО ПРО

Теорема

Если для функций R_1, R_2 и распределения $\mathcal{D}$ задача получения $R_2(x, y)$ по известным xP, yP, xyP и $R_1(x, y)$ является (ε, T)-стойкой, то функция $VKO(x, y, UKM)$ является $(\varepsilon, T - T_{GH} - T_m)$-стойкой в Модели 1, где T_{GH} — время вычисления хэш-функции GH, а T_m — время умножения элемента группы E точек эллиптической кривой на число.

Результат данной теоремы говорит о том, что решение задачи обращения VKO не легче задачи восстановления ключа по журналу протокола Диффи–Хеллмана.

КРИПТО ПРО

Теорема

Если для функций R_1, R_2 и распределения $\mathcal{D}$ задача получения $R_2(x, y)$ по известным xP, yP, xyP и $R_1(x, y)$ является (ε, T)-стойкой, то функция $VKO(x, y, UKM)$ является $(\varepsilon, T - T_{GH} - T_m)$-стойкой в Модели 1, где T_{GH} — время вычисления хэш-функции GH, а T_m — время умножения элемента группы E точек эллиптической кривой на число.

Результат данной теоремы говорит о том, что решение задачи обращения VKO не легче задачи восстановления ключа по журналу протокола Диффи–Хеллмана.

КРИПТО ПРО

Модель 2

Атака: противнику известны xP, yP и UKM.

Угроза: противник находит VKO(x, y, UKM).

КРИПТО ПРО

Обоснование (идейно)

- Пусть есть противник A, который по открытым ключам xP и yP, может эффективно угадывать GH(xyP) с вероятностью p, которая не является пренебрежимо малой.
- Пусть при этом вероятность коллизии у хэш-функции GH на точках кривой E является пренебрежимо малой.
  - В таком случае противник A может решать распознавательную задачу Диффи–Хеллмана (Decisional Diffie–Hellman, DDH, — по xP и yP отличить xyP от Q для случайной точки Q) следующим образом: противник отличает xyP от случайной точки по значению хэш-функции, которую для xyP он может угадывать эффективно и с вероятностью, которая не является пренебрежимо малой.
  - Это противоречит предположению о трудноразрешимости распознавательной задачи Диффи–Хеллмана.

КРИПТО ПРО

Обоснование (идейно) — продолжение

- Предположим теперь, что вероятность коллизии у хэш-функции GH на точках кривой E не является пренебрежимо малой.
  - Тогда существует эффективно конструируемое множество, на котором вероятность коллизии хэш-функции не является пренебрежимо малой.
  - Это позволяет эффективным образом находить коллизию для функции GH.

КРИПТО ПРО

Теорема

Если задача DDH является (ε, T)-стойкой, то функция $VKO(x, y, UKM)$ является $(\varepsilon + s, T - T_{GH} - T_m)$-стойкой в Модели 2, где T_{GH} — время вычисления хэш-функции GH, а T_m — время умножения элемента группы E точек эллиптической кривой на число, а $s = \max\limits_{h \in V_t} \Pr\limits_{Q \in_R E}[GH(Q) = h]$.

Теорема

Если $s = \max\limits_{h \in GH(E)} \Pr\limits_{P \in_R E}[GH(P) = h]$, то для любого $c \in \mathbb{N}$ существует алгоритм, работающий за время T, где $T = \frac{c}{s} + 1$, использующий память объема T, который находит коллизию для хэш-функции GH с вероятностью не меньше $1 - (c + 1) \cdot e^{-c}$.

КРИПТО ПРО

Теорема

Если задача DDH является (ε, T)-стойкой, то функция $VKO(x, y, UKM)$ является $(\varepsilon + s, T - T_{GH} - T_m)$-стойкой в Модели 2, где T_{GH} — время вычисления хэш-функции GH, а T_m — время умножения элемента группы E точек эллиптической кривой на число, а $s = \max\limits_{h \in V_t} \Pr\limits_{Q \in_R E}[GH(Q) = h]$.

Теорема

Если $s = \max\limits_{h \in GH(E)} \Pr\limits_{P \in_R E}[GH(P) = h]$, то для любого $c \in \mathbb{N}$ существует алгоритм, работающий за время T, где $T = \frac{c}{s} + 1$, использующий память объема T, который находит коллизию для хэш-функции GH с вероятностью не меньше $1 - (c + 1) \cdot e^{-c}$.

КРИПТО ПРО

Модель 2’

Атака: противник знает открытые ключи xP,yP, параметр UKM алгоритма VKO_t и некоторую дополнительную информацию $R_3(xyP)$ о точке xyP.

Угроза: противник угадывает результат K работы алгоритма VKO_t.

В этом случае разумными представляются следующие требования: сужение группы точек эллиптической кривой, индуцируемое имеющейся у противника дополнительной информацией $R_3(xyP)$ является эффективно конструируемым, а распознавательная задача Диффи-Хеллмана на этом сужении является трудноразрешимой. В таком случае обоснование стойкости алгоритма VKO_t в модели 2 строится аналогично обоснованию для модели 2 (эффективно конструируемым множеством с высокой вероятностью коллизии для хэш-функции является сужение группы точек эллиптической кривой, индуцируемое дополнительной информацией $R_3(xyP)$).

КРИПТО ПРО

Модель 2'

Атака: противник знает открытые ключи xP,yP, параметр UKM алгоритма VKO_t и некоторую дополнительную информацию $R_3(xyP)$ о точке xyP.
Угроза: противник угадывает результат K работы алгоритма VKO_t.

В этом случае разумными представляются следующие требования: сужение группы точек эллиптической кривой, индуцируемое имеющейся у противника дополнительной информацией $R_3(xyP)$ является эффективно конструируемым, а распознавательная задача Диффи–Хеллмана на этом сужении является труднорешаемой. В таком случае обоснование стойкости алгоритма VKO_t в модели 2' строится аналогично обоснованию для модели 2 (эффективно конструируемым множеством с высокой вероятностью коллизии для хэш-функции является сужение группы точек эллиптической кривой, индуцируемое дополнительной информацией $R_3(xyP)$).

Для каждого из рассматриваемых алгоритмов вопрос стойкости рассмотрен в наиболее сильных моделях противника.

При обоснованиях эксплуатировался стандартный набор предположений о стойкости задач Диффи–Хеллмана в группе точек эллиптических кривых простого порядка и криптографических свойствах введенного стандартом ГОСТ Р 34.11-2012 алгоритма хэширования.

Для всех рассмотренных сопутствующих алгоритмов сделан вывод об их стойкости в соответствующих моделях противника.

Для каждого из рассматриваемых алгоритмов вопрос стойкости рассмотрен в наиболее сильных моделях противника.

При обоснованиях эксплуатировался стандартный набор предположений о стойкости задач Диффи–Хеллмана в группе точек эллиптических кривых простого порядка и криптографических свойствах введенного стандартом ГОСТ Р 34.11-2012 алгоритма хэширования.

Для всех рассмотренных сопутствующих алгоритмов сделан вывод об их стойкости в соответствующих моделях противника.

Для каждого из рассматриваемых алгоритмов вопрос стойкости рассмотрен в наиболее сильных моделях противника.

При обоснованиях эксплуатировался стандартный набор предположений о стойкости задач Диффи–Хеллмана в группе точек эллиптических кривых простого порядка и криптографических свойствах введенного стандартом ГОСТ Р 34.11-2012 алгоритма хэширования.

Для всех рассмотренных сопутствующих алгоритмов сделан вывод об их стойкости в соответствующих моделях противника.

КРИПТО ПРО

Спасибо за внимание!

Выводы

Обоснование стойкости рассматриваемых схем в описанных выше соответствующих моделях противника основывается на следующих свойствах алгоритмов, используемых в качестве базовых.

Свойство 1

Псевдослучайность хэш-функции ГОСТ Р 34.11-2012. Функция GH' является $(\varepsilon_2, T, q, 1)$-стойкой псевдослучайной функцией.

Свойство 2

Стойкость хэш-функции ГОСТ Р 34.11-2012 к атакам со связанными ключами. Пара функций (GH, GH') является (ε_3, T, q, n)-стойкой против атаки с (r', r'')-связанными ключами.

КРИПТО ПРО

Свойство 3

Стойкость хэш-функции ГОСТ Р 34.11-2012 к коллизиям в слабом смысле. Функция GH является (ε_1, T, q, n)-стойкой к коллизиям.

Свойство 3'

Трудоемкость построения коллизий для хэш-функции ГОСТ Р 34.11-2012. Ни один из существующих методов криптографического анализа хэш-функций не позволяет строить коллизии для функции ГОСТ Р 34.11-2012 эффективнее универсальных методов.

Отметим, что свойство 3' предъявляется как представляющийся единственно возможным невырожденный вариант свойства 3, спроецированный на свойства бесключевой функции GH, используемой в алгоритме VKO.