УДК 004.056.001.891 (045)
З 988.0-018.4

Харченко В. С. д-р техн. наук
Халин М. Б.

МНОГОВЕРСИОННОСТЬ ДЛЯ ОБЕСПЕЧЕНИЯ КОНФИДЕНЦИАЛЬНОСТИ И ЦЕЛОСТНОСТИ ИНФОРМАЦИИ: МОДЕЛИ И МЕТОДЫ

Национальный аэрокосмический университет им. Н. Е. Жуковского «ХАИ»

Анализируется возможность применения принципа многоверсионности для обеспечения конфиденциальности и целостности информации. Производится теоретический анализ диверсной хэш-функции.

Введение. Постановка задачи. Растущие потери из-за несанкционированных вторжений в компьютерные системы [1] являются одним из наиболее серьезных вызовов последних десятилетий. Обеспечение безопасности стало важнейшей задачей современных критических и бизнес-критических приложений. Ее решение требует системного анализа и системных мер. В этом контексте целесообразно рассматривать безопасность (*security*) как составляющую более общего свойства – гарантоспособности (*dependability*) [2]. Данный термин соответствует надежности в широком смысле слова, когда, кроме классических составляющих – безотказности, ремонтопригодности, долговечности и сохраняемости, в нее включаются также безопасность в смысле security и в смысле safety, а иногда и живучесть.

При этом отказоустойчивость (прогнозирование, обнаружение, парирование отказов) рассматривается как основное средство обеспечения этих свойств. В современных компьютерных системах критического применения (АСУ АЭС, аэрокосмических комплексах и др.) для создания архитектур, устойчивых к отказам как аппаратных, так и программных средств, применяется принцип диверсности или многоверсионности [3]. Его сущность состоит в реализации одних и тех же функций в избыточных компьютерных системах с помощью различных программно-аппаратных средств. Появились работы, в которых признается перспективность его использования для задач защиты информации [4-6].

Цель данной статьи – анализ возможности и оценка ожидаемого эффекта применения принципа многоверсионности для реализации основных услуг систем защиты – конфиденциальности и целостности информации.

Конфиденциальность. Пусть существует сообщение M, смысловое содержание которого необходимо скрыть. Применив к сообщению M [4] n различных алгоритмов криптопреобразования (шифров), получим n криптограмм:

$$F_1(M)\to C_1,\ F_2(M)\to C_2 \ldots F_n(M)\to C_n. \quad (1)$$

Проведем сравнительный анализ многоверсионного шифрования с общепринятым применением одного шифра для одного сообщения:

$$F(M)\to C. \quad (2)$$

1. Если сложность криптоанализа одного из шифров меньше, чем сложности криптоанализа других шифров $I_j=\min\{I_k\},\ j\in[1,n], k=\overline{[1,n]}$, то тогда криптоаналитику необходимо совершить успешную атаку только лишь на C_j для того, чтобы раскрыть содержание M.

2. Если $I_1=I_2=\ldots=I_n$, то повышается вероятность появления новой аналитической атаки на один из шифров, что ведет к ситуации, описанной в предыдущем пункте.

3. Если какой-либо из шифров имеет меньшую вычислительную сложность прямого преобразования

$$I_j^{ПР} = \min\{I_k^{ПР}\},\ j \in [1,n], k = \overline{[1,n]}, \quad (3)$$

то это его неопровержимое преимущество сводится на нет шифром, который имеет максимальную вычислительную сложность

$$I_l^{ПР} = \max\{I_k^{ПР}\},\ l \in [1,n], k = \overline{[1,n]}, \quad (4)$$

так как после криптографических преобразований должна производиться синхронизация многоверсионной системы шифрования.

4. Увеличение количества ключевых систем повлечет за собой рост затрат на их генерацию, доставку, хранение, архивирование:

$$C_n \approx nC^1, \quad (5)$$

где C_n – затраты в многоверсионной системе, C_1 – затраты при использовании одного шифра.

5. При многоверсионном методе шифрования общая длина криптограмм

$$L_{\sum} \approx nL_C, \quad (6)$$

где L_C – длина криптограммы C, что приводит к:

а) уменьшению пропускной способности канала, если сообщение M необходимо передать по сети;

б) существенному увеличению емкости носителя, на котором необходимо сохранить криптограммы $C_i, i = \overline{[1,n]}$.

Аутентификация, целостность. Пусть существует сообщение M [4] целостность и авторство, которого необходимо сохранить. Для реализации этих функций СЗИ с применением принципа многоверсионности вырабатывается n цифровых подписей (ЦП) на основе сообщения M:

$$F_i^{ЦП}(M) = D_i, i = \overline{[1,n]}. \quad (7)$$

Проведем сравнительный анализ многоверсионной ЦП с общепринятым применением одной ЦП для одного сообщения:

$$F^{ЦП}(M) = D. \quad (8)$$

1) Если криптоаналитик сумеет провести атаку на $\left\lfloor \frac{n-1}{2} \right\rfloor$ ЦП, то возможно определить на какие именно алгоритмы ЦП производилась атака, а также выявить истинное сообщение. В случае (8) успешная атака на одну ЦП приведет к тому, что злоумышленник сможет генерировать собственные сообщения с произвольным авторством.

2) Так как при генерации ЦП используется хэш-функция, то большее множество сообщений M отображается в меньшее множество ЦП D. Таким образом, одной ЦП могут соответствовать несколько сообщений из подмножества $M' \in M$ (т.е. возникают так называемые коллизии), что приводит к возможности подмены криптоаналитиком сообщения M без изменения ЦП D и без знания секретного ключа ЦП при использовании стандартной системы ЦП (8), т. е.:

$$F^{ЦП}(M_i) = D, \forall M_i \in M'. \quad (9)$$

При использовании многоверсионной ЦП (12.7) существуют множества тождественных сообщений для каждого из алгоритмов ЦП:

$$M'_1 \in M, M'_2 \in M, \ldots, M'_n \in M,$$

но подмножество сообщений, которое тождественно относительно всех ЦП, лежит на пересечении всех подмножеств тождественных сообщений каждой из ЦП:

$$M'_{ОБ} = M'_1 \cap M'_2 \cap \ldots \cap M'_n. \quad (10)$$

Таким образом, в самом худшем случае

$$M'_{ОБ} = M'_i, i \in \overline{[1,n]}, \quad (11)$$

причем мощность подмножества M'_i является минимальным среди подмножеств тождественных сообщений для всех ЦП, т. е.

$$|M'_i| = \min\{|M'_k|\}, k = \overline{[1,n]}. \quad (12)$$

В самом лучшем случае $M'_{ОБ} = \emptyset$, т. е. не будет существовать ни одного сообщения $M_a \in M$, которое было бы тождественно сообщению $M_b \in M, M_a \neq M_b$, относительно всех ЦП используемых в многоверсионной системе.

3) Если криптоаналитик сумеет провести атаку на $n\text{-}k, k \in [1, n-1]$ ЦП с подменой сообщений, то возможно установить факт НСД с вероятностью достоверности

$$P_{НСД} = 1 - P_{ТП}, \quad (13)$$

где $P_{ТП}$ – вероятность того, что сообщение $M_{a,i}$ с соответствующей ЦП $D_{a,i}$ после случайной помехи в канале преобразуется в сообщение $M_{b,i}$ с соответствующей ЦП $D_{b,i}$. Очевидно, что $P_{ТП} \to 0 \Rightarrow P_{НСД} \to 1$, но для вычисления точных значений данных вероятностей необходим более детальный анализ.

4) Если какая-либо из ЦП имеет меньшую вычислительную сложность, то это её неопровержимое преимущество сводится на нет ЦП, которая имеет максимальную вычислительную сложность, так как после криптографических преобразований должна производиться синхронизация многоверсионной ЦП.

5) Увеличение количества используемых ЦП в СЗИ приведет к такому же увеличению ключевых систем, что, в свою очередь, повлечет за собой рост затрат на их генерацию, доставку, хранение, архивирование:

$$C^{ЦП}{}_n \approx nC^{1,ЦП}, \qquad (14)$$

где $C^{ЦП}{}_n$ – затраты в многоверсионной системе, $C^{1,ЦП}$ – затраты при использовании одной ЦП.

6) При многоверсионном методе ЦП общая длина подписанного сообщения не увеличивается прямопропорционально количеству применяемых алгоритмов ЦП n, т. к. нет необходимости подписывать несколько копий одного и того же сообщения (если оно отправляется только по одному каналу передачи информации). Тогда общая длина подписанного сообщения многоверсионной ЦП:

$$L_{\Sigma}^{ЦП} \approx L_M + nL_{ЦП}, \qquad (15)$$

где L_M – длина сообщения M, $L_{ЦП}$ – длина одной ЦП. Учитывая, что $L_{ЦП} << L_M$ верно следующее:

а) уменьшение пропускной способности канала, если сообщение М необходимо передать по сети носит незначительный характер;

б) увеличение емкости носителя, на котором необходимо сохранить сообщение, подписанное несколькими ЦП, также носит незначительный характер.

Таким образом, применение принципа многоверсионности для повышения эффективности СЗИ носит противоречивый характер. С одной стороны, рассматривая наиболее общий подход применения МВТ к услуге конфиденциальности СЗИ, был выявлен ряд недостатков данного метода, что и следовало ожидать, т.к. параллельное применение различных алгоритмов криптопреобразования к одному сообщению с целью скрытия смысла дает лишь возможность выбора для криптоаналитика средств и метода криптоанализа для выполнения успешной атаки.

С другой стороны, существует особый подвид многоверсионных криптоалгоритмов – комплексные криптоалгоритмы, которые при некоторых условиях являются эффективным решением задачи обеспечения конфиденциальности в СЗИ. В качестве примера такого алгоритма можно привести широко известный *RSA-OAEP*.

Более эффективным является применение принципа МВТ для построения ЦП, с чьей помощью повышается вероятность получения подлинного сообщения. Особого внимания заслуживает возможность понижения вероятности появления коллизий ЦП из-за применения различных алгоритмов хэш-функций. Далее приведена теоретическая оценка понижения мощности множества тождественных сообщений относительно диверсной хэш-функции как составной части диверсной ЦП.

Оценка понижения мощности множества тождественных сообщений относительно диверсной хэш-функции. Необходимо найти вероятность непустого пересечения множеств тождественных сообщений относительно двух хэш-функций H_1 и H_2, имеющих дайджест одинаковой длины, т.е. вероятность выполнения условия $M'_{ОБ} = M'_1 \cap M'_2 \neq \emptyset$. Пусть n – длина сообщений в битах, k – длина дайджестов в битах каждой хэш-функций. Количество тождественных сообщений относительно одного значения хэш-функции будет равным 2^{n-k}. Следовательно, мощность множеств тождественных сообщений относительно хэш-функций H_1 и H_2:

$$|M'_1| = |M'_2| = 2^{n-k}. \qquad (16)$$

Обозначим элементы [2] множеств тождественных сообщений символами

$M'_{i,j}$, где i – номер множества тождественных сообщений, j – номер элемента в этом множестве. Для $M'_{1,1}$ вероятность того, что $M'_{2,1} = M'_{1,1}$, равна $\frac{1}{2^n}$, поэтому вероятность несовпадения $M'_{2,1}$ с $M'_{1,1}$ равна $\left(1-\frac{1}{2^n}\right)$. Если рассмотреть 2^{n-k} случайных значений из множества M'_2, вероятность того, что ни одно из этих значений не совпадает с $M'_{1,1}$, составляет $\left(1-\frac{1}{2^n}\right)^{2^{n-k}}$. Поэтому вероятность по крайней мере одного совпадения с $M'_{1,1}$ равна $1-\left(1-\frac{1}{2^n}\right)^{2^{n-k}}$. Тогда вероятность непустого пересечения $P(M'_{ОБ} \neq \emptyset)$ вычисляется следующим образом:

$$P(M'_{ОБ} \neq \emptyset) = 1-\left(1-\frac{1}{2^n}\right)^{2^{2(n-k)}}. \quad (17)$$

Данное выражение является сложным для вычисления поэтому упростим его используя следующее неравенство:

$$(1-x) \le e^{-x} \quad \text{для всех } x \ge 0. \quad (18)$$

Следует обратить внимание, что так как функция (1-x) является касательной к функции e^{-x} в точке x=0, то при

$$x \to 0 : (1-x) \approx e^{-x}. \quad (19)$$

Тогда формулу (12.17) можно представить в виде:

$$P(M'_{ОБ} \neq \emptyset) > 1-\exp\left(-\frac{2^{2(n-k)}}{2^n}\right), \quad (20)$$

$$P(M'_{ОБ} \neq \emptyset) > 1-\exp\left(-2^{n-2k}\right). \quad (21)$$

Теперь сформулируем следующую задачу: при каком значении k можно получить неравенство

$$P(M'_{ОБ} \neq \emptyset) > 0{,}5. \quad (22)$$

Для этого потребуется, чтобы

$$\frac{1}{2} = 1-\exp\left(-2^{n-2k}\right),$$

$$\ln(2) = 2^{n-2k},$$

$$k = \frac{n-\log_2\left(\frac{1}{\ln(2)}\right)}{2} \approx \frac{n-0{,}53}{2} \approx \frac{n}{2}. \quad (23)$$

Таким образом, если использовать две хэш-функции, имеющих дайджесты длиной 32 бита, то при их применении к сообщению длиной 64 бит будет выполняться неравенство (22). Это значит, что с большой вероятностью будет существовать еще хотя бы одно сообщение (помимо исходного) для которого значения обоих хэш-функций будут такими же, как и для исходного сообщения.

Выводы. Применение принципа многоверсионности для повышения эффективности СЗИ носит противоречивый характер. С одной стороны, рассматривая наиболее общий подход применения МВТ к услуге конфиденциальности СЗИ был выявлен ряд недостатков данного метода, что и следовало ожидать, т. к. параллельное применение различных алгоритмов криптопреобразования к одному сообщению с целью скрытия смысла дает лишь возможность выбора для криптоаналитика средств и метода криптоанализа для выполнения успешной атаки. Однако, с другой стороны, существует особый подвид многоверсионных криптоалгоритмов – комплексные криптоалгоритмы, которые при некоторых условиях являются эффективным решением задачи обеспечения конфиденциальности в СЗИ.

Также эффективным является применение принципа МВТ для построения ЦП, с чьей помощью повышается вероятность получения подлинного сообщения. Особого внимания заслуживает возможность понижения вероятности появления коллизий ЦП из-за применения различных алгоритмов хэш-функций.

Развитием данной работы должно послужить экспериментальное подтверждение или опровержение приведенных выше зависимостей.

Список литературы

1. *S. Garfinkel, G. Spafford.* WebSecurity, Privacy and Commerce, Beijing, Cambridge, Farnham, Köln, Paris, Sebastopol', Taipei, Tokio, 'Reilly, 2002. – 756 p.

2. *A.Avižienis, J-C.Lapri, B.Randel.* Fundamental Concepts of Dependability, UCLA CSD Report no.010028, LAAS Report no.01-145, Newcastle University Report no. CS-TR-739, 2002. – 20 p.

3. *Многоверсионные* системы, технологии и проекты. Под ред. В. С. Харченко, Минобразования и науки Украины, НАКУ «ХАИ»: Харьков, 2003. – 486 с.

4. *Харченко В. С., Халин М. Б.* О многоверсионной цифровой подписи// Материалы МНТК «Информационные технологии в машиностроении»: Харьков, НАКУ «ХАИ», 2003. – С.24.