Математические структуры и моделирование
2003, вып. 11, с. 172–174

УДК 519.68

ЗАЩИТА БАЗЫ ДАННЫХ ФОТОДОКУМЕНТОВ

Д.Н. Лавров, А.В. Мухоморов

In article method protection database is presented. The method is combination of the block algorithm of the cryptooperation and stego-algoritm.

Защита базы данных (БД) — сложная многоплановая задача. Цели защиты могут быть различными. Так, в [2] ставится задача защиты таблицы БД, каждая запись которой состоит из двух полей: «фамилия сотрудника» и «данные о сотруднике». Первое поле является индексным, второе поле данных. Выбирается однонаправленная хэш-функция и симметричный алгоритм шифрования. Поле данных шифруется с помощью симметричного алгоритма, использующего в качестве ключа данные индексного поля. Не зная данных индексного поля, невозможно получить данные о сотруднике.

Такая защита не позволяет получить сразу данные о всех сотрудниках. Поиск по маске также невозможен.

Эта защита не совершена [4], так как может быть вскрыта с помощью «грубого» взлома [1].

Проблемы защиты БД путем прямого шифрования полей данных обсуждаются в [3].

Целью данной работы является разработка системы защиты БД фотодокументов, которая, с одной стороны, позволяла бы получить полный доступ к базе, включая контекстный поиск, а с другой – без специального программного обеспечения невозможно было бы использовать и преобразовывать данные.

База данных фотодокументов состоит из записей, содержащих поля: «фонд», «опись», «дело», «лист», «текст», «фото». В поле «текст» хранится текст документа, в поле «фото» — отсканированное фотоизображение документа. Два поля «текст» и «фото» хранятся в различных таблицах БД. Цели этого разнесения будут прояснены ниже.

Так же как и в [2], выбирается однонаправленная хэш-функция h и симметричный алгоритм шифрования. Алгоритм зашифрования на ключе K обозначим $E_K(\cdot)$, расшифрования – $D_K(\cdot)$. В программной реализации использовались хэш-функция MD5 и алгоритм Blowfish. Введем обозначения: T – данные поля «текст», P – данные поля «фото» текущей записи. Алгоритм защиты БД фотодокументов состоит в следующем:

E-mail: petroff@mail.ru
Омский государственный университет

1. Вычисляем ключи шифрования $h(T)$ и $h(P)$.
2. С помощью полученных ключей «перекрестно» шифруем поля «текст» и «фото», получаем $E_{h(P)}(T)$ и $E_{h(T)}(P)$.
3. Размещаем ключи $h(T)$ и $h(P)$ в самих полях с помощью алгоритма вставок в случайные позиции, определенные специальным ключом K. Обозначим эту функцию $S_K(\cdot,\cdot)$ (на первой позиции данные поля, на второй ключ), тогда преобразованные данные описываются композицией отображений:
$$S_K(E_{h(P)}(T), h(T)) \quad \text{и} \quad S_K(E_{h(T)}(P), h(P)).$$
4. Преобразованные данные помещаются в поля «текст» и «фото».

Ключ K представляет собой номера байт, в которых хранятся ключи $h(T)$ и $h(P)$ в контейнерах $E_{h(T)}(P)$ и $E_{h(P)}(T)$ соответственно. Ключ K должен быть различным для различных записей. Этого можно достичь с помощью криптостойкого генератора псевдослучайной последовательности.

Алгоритм S размещения ключей шифрования в данных является по сути алгоритмом стеганографии, так как в качестве сообщения выступает ключ, а в качестве контейнера – данные полей записи.

Доступ к данным осуществляется по следующему алгоритму:
1. Вынимаем ключи $h(T)$ и $h(P)$ из поля текущей записи.
2. С помощью полученных ключей расшифровываем поля «текст» и «фото», получаем $P = D_{h(T)}(E_{h(T)}(P))$ и $T = D_{h(P)}(E_{h(P)}(T))$.

Если злоумышленник получит доступ только к одной таблице БД, в которой хранятся данные $S_K(E_{h(P)}(T), h(T))$, то прочесть данные из нее он не сможет, так как ему неизвестен ключ $h(P)$. Единственное, что он сможет получить, – это хэш $h(T)$, да и то если известен стегоключ K. В этом случае грубый взлом неэффективен.

Проанализируем ситуацию, когда злоумышленник получил все таблицы БД. Рассматривая запись БД, злоумышленник получает в свое распоряжение пару $(S_K(E_{h(P)}(T), h(T)); S_K(E_{h(T)}(P), h(P)))$. Оценим трудоемкость атаки грубой силы на стегоалгоритм. Пусть N – размер контейнера в байтах, n – размер ключа в байтах. Тогда вариантов размещения ключа в контейнере будет

$$A^n_{N+n} = \frac{(N+n)!}{n!} = (n+1)\cdot\ldots\cdot(N+n) \geq (n+1)^N.$$

Следовательно, время, потраченное на вскрытие одной записи БД, больше, чем

$$(n+1)^N \cdot t,$$

где t – время проверки подлинности одного ключа.

Пусть для примера $n = 256$ бит$=$ 32 байта (типичная длина ключа), $N = 2400$ байт (соответствует странице текста), $t = 10^{-3}$ сек., тогда время, потраченное на вскрытие одной записи, будет составлять $33^{2400}\cdot 10^{-3} \approx 3\cdot 10^{3657}$ секунд.

Наиболее опасный вариант – это когда атакующий получает не только таблицы базы, но и программное обеспечение доступа и работы с БД. При анализе выполнимого кода программного обеспечения возможно отследить работу стегоалгоритма и получить ключ K. Для предотвращения подобной атаки необходима защита программного кода. Возможны следующие варианты защиты кода: защита от отладки (препятствует работе отладчиков типа SoftIce), защита от дизассемблированния (код стегоалгоритма шифруется при хранении в EXE-файле); логические ловушки для отладчиков; привязка программного кода к аппаратной части компьютера.

Важно, чтобы стоимость взлома сравнялась со стоимостью данных и программного обеспечения.

Наиболее уязвимым местом данной системы защиты является защита программного обеспечения. Но удовлетворительного решения в мире компьютерных технологий неизвестно до сих пор.

В заключение отметим, что платой за представленную защиту БД является, как показывают эксперименты, увеличенное в среднем вдвое время обращения к записям базы.

Литература

1. Blakley G.R., Meadows C. *A Database Encryption Scheme which Allows the Computation of Statistics Using Encrypted Data* // Proceedings of the 1985 Symposium on Security and Privacy, IEEE Computer Society. 1985. Apr. P. 116-122.
2. Feigenbaum J., Liberman M.Y., Grosse E., Reeds J.A. *Cryptographic Protection of Membership Lists* // Newsletter of the International Association of Cryptologic Research. 1992. V. 9. P. 16-20.
3. Епанчинцева О.Л., Ворошилов В.В. Шифрование данных в DBF-файлах // Математические структуры и моделирование. Сб. науч. трудов. / Под ред. А.К.Гуца. Омск: ОмГУ, 2000. №6. С.139-142.
4. Шнаейр Б. *Прикладная криптография.* М.: Триумф. 2002. 816 с.