БЛОЧНЫЕ ШИФРЫ И ИХ КРИПТОАНАЛИЗ

А. И. Пестунов
Институт вычислительных технологий СО РАН, Новосибирск, Россия
e-mail: an24@ngs.ru

Nowadays block ciphers became the most important way for data protection, therefore there are a number of projects devoted to their investigation Specialists from all over the world are involved in their design and analysis. This paper provides a survey of design principles and methods of checking the security of contemporary block ciphers.

Введение

Из-за широкого распространения сети Интернет передавать информацию на большие расстояния стало значительно проще, чем, скажем, несколько десятилений назад, но существенно обострилась проблема ее защиты на пути следования от отправителя к получателю. Конечно, посланный обычным человеком e-mail вряд ли представляет интерес для посторонних лиц, но, например, правительственное сообщение или карточка больного, принадлежащая известному человеку, несомненно, вызывают желание ознакомиться с ними у многих людей. Стоит заметить, что целью злоумышленника может быть не просто получение данных, содержащихся в сообщении, но и их фальсификация с тем, чтобы адресат получил ложные сведения, предполагая тем не менее, что они — верные.

Для того чтобы злоумышленник не смог извлечь пользу из перехвата информации или получатель смог определить, что сообщение фальсифицировано, как правило, используются *схемы с секретным* (закрытым) *ключом*. По большому счету, они уже применялись в древности, но формализованы были в 1949 г. Клодом Шенноном (Claude Shannon), именно со времени его публикации [1] *криптография*, которая ранее называлась *тайнописью*, считается оформившейся как наука.

Схема с секретным ключом (рис. 1) подразумевает наличие двух каналов связи: “дешевого” (Интернет, телефон) и “дорогого” (личная встреча, курьерская почта). Дешевизна первого из них может заключаться в его материальной стоимости, доступности в любой момент или высокой скорости передачи сообщений. Дорогой канал не предоставляет таких удобств, но он, в отличие от дешевого, защищен от доступа к нему посторонних. Для безопасного общения отправитель и получатель информации разово обмениваются секретным ключом с помощью защищенного канала, а затем передают сообщения, зашифрованные некоторым алгоритмом с полученным ключом, по незащищенному каналу.

В 1976 г. Уидфилд Диффи (Whitfield Diffie) и Мартин Хеллман (Martin Hellman) [2] предложили *схему с открытым ключом*, которая позволяет обмениваться данными без

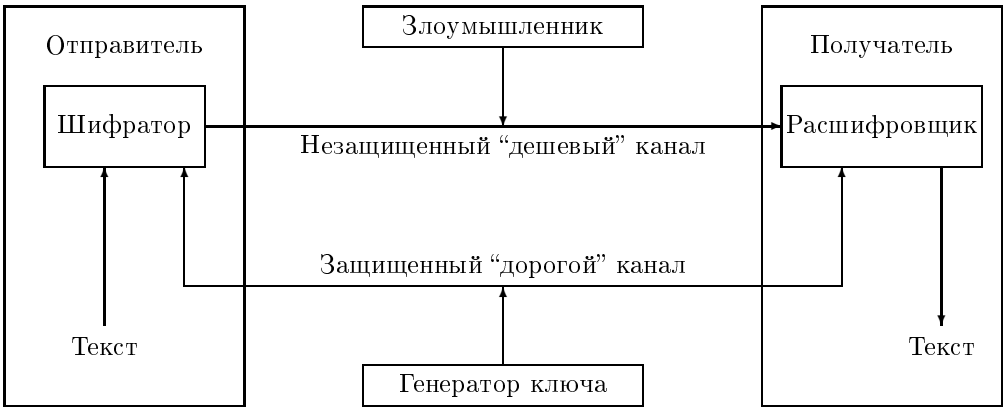


Рис. 1. Схема с секретным ключом

наличия дорогого защищенного канала. Скорость работы подобного рода схем мала, и они обычно применяются в комбинации со схемами с секретным ключом, выступая в роли дорогого канала: абоненты формируют секретный ключ с помощью схемы Диффи — Хелмана (или аналогичной ей) и осуществляют обмен сообщениями с помощью схемы Шеннона.

В качестве шифрующего алгоритма в схеме с секретным ключом чаще всего используются блочные или поточные шифры. Поточный шифр представляет собой зависящий от ключа генератор псевдослучайных битов, которые XORируются с битами открытого текста, формируя шифртекст. Для расшифровки сообщения необходимо произвести те же операции, только с шифртекстом. Несмотря на кажущуюся простоту поточного шифра, сами алгоритмы генерации битов не имеют стройной, хорошо изученной теории их построения и анализа, в то время как блочные шифры изучены существенно лучше.

В настоящей статье дан обзор развития блочного шифрования в период последней четверти XX в. и до настоящего времени. Представлена базовая информация, касающаяся блочных шифров, показаны основные варианты их дизайна и анализа. Завершается статья кратким обзором международных конкурсов, направленных на исследование блочных шифров.

1. Общие сведения о блочных шифрах

Типичный блочный шифр преобразует открытый текст, представленный в виде последовательности битов (нулей и единиц), по блокам фиксированной длины, чаще всего 64 или 128 бит. Секретный ключ также представляет собой битовую последовательность обычно длиной 128 или 256, из которой с помощью определенного для каждого шифра алгоритма получается набор из R *подключей*. Процедура шифрования заключается в R-кратном выполнении некоторой относительно простой функции, которая зависит от своего подключа и называется *раундом* шифрования. Чаще шифр состоит из одинаковых раундов, но иногда они различаются. Чем больше раундов выполнено, тем надежней, но медленней становится шифр, поэтому создатель определяет такое количество раундов, которое обеспечит как безопасность, так и быстродействие алгоритма. Для удобства реализации, особенно на физическом уровне, блочные шифры проектируются так, чтобы процедуры шифрования и дешифрирования были похожи. Наиболее удачными в этом отношении выглядят те шифры, в которых эти две процедуры разли-

чаются только порядком подачи подключей, т. е. для расшифровки требуется обратный порядок подключей с R-го по 1-й.

Главные свойства шифра — это пропускная способность (скорость шифрования) и надежность. Скорость измеряется числом тактов на шифрование одного блока или просто физическим временем, если измерения проводятся на аналогичных процессорах. Существенно сложнее обстоит дело с безопасностью. Надежным шифром интуитивно можно считать тот, который невозможно взломать, или вскрыть. Но это понятие требует уточнения. Для того чтобы внести определенность, еще в 1883 г. Август Керхгоффс (Auguste Kerckhoffs) предложил считать, что злоумышленник знает о криптосистеме все, кроме секретного ключа, и главной его целью является разработка метода, позволяющего отыскать ключ, используя какие-то эксперименты с шифром[1]. Алгоритм такого сорта называется *атакой*.

Отсюда следует, что наиболее естественным способом изучения надежности шифра является попытка построить атаку на данный шифр, причем делать это должен прежде всего его создатель, и, если обнаружится, что шифр может быть успешно атакован, то он слабый и требует усовершенствования. Чтобы сравнивать разного рода атаки в едином ключе, в большинстве современных работ определяются требования для их реализации: число зашифрованных блоков текста, память и число операций шифрования.

Любой шифр может быть вскрыт с помощью полного перебора ключей. Злоумышленник перехватывает одну пару (или несколько пар) блоков — "текст" и "шифртекст", а затем, шифруя "текст" всеми возможными ключами, ищет тот из них, который приведет к шифрованию "текста" в "шифртекст". Очевидным достоинством данного метода является то, что таким образом вскрывается любой шифр. Главный недостаток — необходимость перебирать все возможные ключи. Так, например, если длина ключа составляет 128 бит, то такой метод требует в среднем 2^{127} операций шифрования. Отсюда следует, что длина ключа должна быть достаточной, чтобы исключить подобного рода угрозы, причем и в обозримом будущем, поскольку мощности ЭВМ растут.

Считается, что достаточная безопасность достигается при как минимум 80-битном ключе. Заметим, что перебор, кажущийся нереальным для одного процессора, может достаточно быстро осуществиться, если, например, подключатся все граждане страны, желающие вскрыть шифровку неприятеля. Взломом шифра считается любая атака, работающая быстрее полного перебора ключей, особый интерес, конечно, представляют те из них, которые не просто быстрее полного перебора, но и осуществимы на практике.

Первым блочным шифром стоит признать алгоритм Люцифер (Lucifer), опубликованный в 1973 г. [3], который разработан Хорстом Фейстелем, а через четыре года Национальным бюро стандартов США был опубликован усовершенствованный вариант алгоритма Люцифера — шифр DES — Data Encryption Standard (стандарт шифрования данных), который играл ключевую роль в защите информации на протяжении последней четверти XX в. [4]. DES принимал 56-битный ключ, но с развитием вычислительной техники перебор размера 2^{56} становился все более и более реальным, компанией RSA был даже запущен Интернет-проект, в рамках которого люди по всему миру осуществляли подобный перебор, и примерно за два года ключ был найден.

Для того чтобы увеличить длину ключа, были предложены различные усовершенствования DES, самый успешный из них — 3-DES, который за небольшими изменениями является троекратным шифрованием с помощью DES и имеет 168-битный ключ. Без-

[1] В широком смысле взломом шифра считается любой найденный недостаток, который может хотя бы отдаленно намекнуть на несовершенство шифра.

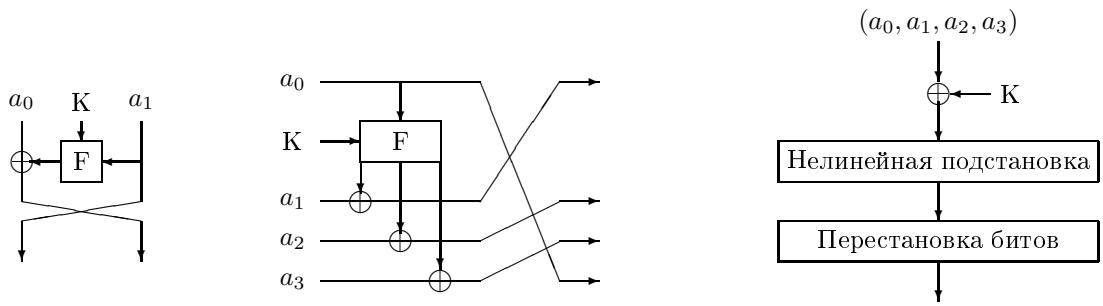
опасность этого алгоритма не вызывает сомнений, но скорость шифрования и относительная сложность реализации были неприемлемы для ряда приложений и устройств. Очевидная необходимость обновления стандарта шифрования произвела своего рода бум, в результате которого появились новые алгоритмы, претендовавшие на потенциально вакантное место, а в 1997–2000 гг. Национальный институт стандартов и технологий США провел конкурс шифров, победитель которого получил название AES (Advanced Encryption Standard) [5] и был принят в качестве стандарта.

2. Дизайн блочного шифра

Блочные шифры конструируются двумя наиболее распространенными способами: в виде сети Фейстеля и подстановочно-перестановочной сети. Сеть Фейстеля получила свое название по имени Хорста Фейстеля — создателя первого блочного шифра Люцифер. Раунды этого типа шифруют блок текста следующим образом: блок вначале разбивается на две равные части, затем правая из них преобразуется функцией F, зависящей от подключа, и XORируется с левой частью, после этого части меняются местами. Сеть Фейстеля очень хороша тем, что шифрование отличается от дешифрирования только порядком подачи подключей, а раундовая функция может быть практически любой. Главный недостаток такого рода шифров заключается в том, что за один раунд изменяется только одна половина блока. Позднее появились различные модификации сети Фейстеля (рис. 2), что связано в основном с тем, что классическая сеть Фейстеля ориентирована на 64-битные блоки, а если преобразовывать текст по 128-битным блокам, то возникают проблемы, как как платформы на большинстве современных ЭВМ 32-битные. Вследствие этого блок необходимо разбивать не на две, а на четыре части. Одна из модификаций сети Фейстеля приведена на рис. 2.

Подстановочно-перестановочная сеть позволяет преобразовывать весь блок за один раунд (рис. 2). Один раунд такого шифра состоит из трех обратимых преобразований: сложение с подключом, нелинейная подстановка и перестановка битов. Эта структура хороша своей относительной прозрачностью и возможностью настраивать отдельные компоненты так, чтобы они имели нужные свойства, например устойчивость к определенным атакам.

Шифр DES представляет собой сеть Фейстеля, где в качестве функции F используется небольшая подстановочно-перестановочная сеть. Победитель конкурса AES шифр



Сеть Фейстеля — Модифицированная сеть Фейстеля — Подстановочно-перестановочная сеть

Рис. 2. Типы раундов блочного шифра

Rijndael [6] полностью является подстановочно-перестановочной сетью с двумя типами перестановок. Модифицированными сетями Фейстеля являются финалисты конкурса AES: Mars, RC6 и Twofish.

3. Криптоанализ блочного шифра

Как было сказано выше, секретный ключ напрямую не участвует в шифровании: с его помощью формируется массив подключей, от которых зависят раундовые функции. По этой причине большинство атак ориентировано именно на нахождение подключей, а не секретного ключа.

В рамках принципа Керхгоффса можно определить несколько сценариев атак на блочные шифры, каждый из которых накладывает некоторые ограничения на возможности злоумышленника. *Атака по известному шифртексту* подразумевает, что он может только перехватывать зашифрованные сообщения. При *атаке по известному открытому тексту* у злоумышленника есть возможность знать не только шифртекст, но и соответствующий ему открытый текст. *Атака по выбранному открытому тексту* позволяет взломщику шифровать любые сообщения по своему усмотрению и исследовать шифртекст. Надо сказать, что этот с первого взгляда нереальный сценарий вполне возможен. Рассмотрим пример, связанный с электронной почтой. Предположим, что злоумышленник желает узнать, каким ключом некий человек шифрует свои письма. Он отправляет этому человеку письма с какими-нибудь вопросами, на которые он уже знает ответы, а затем перехватывает зашифрованные письма, идущие по незащищенному каналу связи. В итоге, подбирая специальным образом вопросы, злоумышленник узнает, какой шифртекст представляют известные ему ответы после шифрования.

3.1. Дифференциальный и линейный криптоанализ

Дифференциальный криптоанализ [7], предложенный Эли Бихамом (Eli Biham) и Ади Шамиром (Adi Shamir) в 1992 г., стал первым методом, который позволил разработать алгоритм поиска всех подключей для DES быстрее полного перебора ключей. В том же году Мицуру Мацуи (Mitsuru Matsui) предложил линейный криптоанализ [8], что позволило находить подключи еще быстрее.

Основной объект, который исследуется в дифференциальном криптоанализе, — это пары блоков текста A и B с определенной разностью (difference), как правило, эта разность определяется как исключающее или (xor), т. е. $A \oplus B$. Если мы не располагаем информацией о том, как связаны входная (между блоками открытого текста) и выходная (между блоками шифртекста) разности, то все выходные разности равновероятны, но если нам удалось установить, что некоторая входная разность Δ_{inp} вызывает некоторую выходную разность Δ_{out} с вероятностью, большей, чем остальные, то это может быть использовано для поиска подключей шифра. Пара (Δ_{inp} и Δ_{out}) называется *дифференциалом*, а совокупность всех дифференциалов на различных раундах называется *характеристикой*.

Алгоритм поиска подключей поясним на примере. Рассмотрим шифр, имеющий 64-битный блок, 128-битный секретный ключ, состоящий из $R = 16$ раундов и на всех раундах использующий 32-битные подключи. Пусть пара блоков A и B имеет разность Δ_{inp}. Если шифр совершенный, то вероятность того, что разность пары блоков шифр-

текста равна Δ_{out}, составляет 2^{-64}, т. е. все выходные разности равновероятны. Предположим, нами установлено, что если такую пару зашифровать с помощью $R-1$ раундов, то выходная разность равна Δ_{out} с вероятностью 2^{-20}. Такой недостаток шифра может быть использован для поиска подключа R-го раунда.

Атака реализуется в несколько шагов. Возьмем 2^{20} пар блоков с разностью Δ_{inp}, зашифруем их с помощью R раундов и сохраним в памяти. Перебираем все возможные подключи R-го раунда и для каждого из них выполняем следующие действия: расшифровываем на один раунд сохраненные в памяти пары шифртекста и проверяем равенство их разности Δ_{out}. Если ни одна из пар не приводит к такой разности, то отбрасываем этот подключ и переходим к следующему. Если такая разность появилась хотя бы раз, то подключ — верный.

Будем говорить, что атака завершилась *успехом*, если верный подключ раунда R найден. Это означает, что все неправильные подключи отброшены, а правильный остался. Оценим вероятность успеха для нашей атаки. Пара, расшифрованная с помощью неправильного подключа, не удовлетворяет дифференциалу с вероятностью 2^{-20}, поэтому вероятность появляения любой разности (в том числе и Δ_{out}) равна 2^{-64}. Вероятность появления такой разности среди 2^{20} пар равна 2^{-44}. Вероятность, что хотя бы один из $2^{32}-1$ неправильных подключей приведет к такой разности, равняется 2^{-12}, и все неправильные подключи будут отброшены с вероятностью $1-2^{-12}$.

Если пара расшифрована с помощью правильного подключа, то выходная разность равна Δ_{out} с вероятностью 2^{-20}, а, используя распределение Пуассона, можно заключить, что среди 2^{20} пар одна будет с нужной разностью — с вероятностью 0.63. Таким образом, вероятность успеха нашей атаки равна примерно 0.63, она может быть увеличена, если увеличить число блоков текста.

Теперь подсчитаем сложность атаки. Необходимо 2^{21} зашифрованных выбранных блоков открытого текста, 2^{24} байт памяти, чтобы их хранить, и 2^{53} операций однораундового дешифрирования (эквивалентно 2^{49} полнораундовых шифрований), чтобы для каждого проверяемого подключа (из 2^{32}) обработать все блоки текста. Метод полного перебора ключей требует в среднем 2^{127} операций шифрования.

После описания исходного варианта дифференциального криптоанализа стали появляться различные его модификации: усеченные (truncated) дифференциалы [9], атака бумерангом (boomerang attack) [10], нереальные (impossible) дифференциалы [11]. На данный момент дифференциальный криптоанализ вместе со всеми своими модификациями является бесспорным лидером по числу успешных попыток вскрытия блочных шифров.

В линейном криптоанализе изучаются не блоки с фиксированной разностью, а некоторые линейные выражения, включающие биты подключа, текста и шифртекста. Такое выражение будет верно с вероятностью $1/2$ для совершенного шифра, но если будет найдено отклонение от $1/2$, то это может быть использовано для поиска подключа достаточно схожим с дифференциальным криптоанализом образом, но в отличие от дифференциального криптоанализа линейный является атакой по известному, а не выбранному открытому тексту.

3.2. Интегральный криптоанализ и другие виды атак

После успешного применения дифференциального и линейного криптоанализа к ряду шифров новые варианты шифров стали конструироваться такими, чтобы быть устой-

чивыми к этим атакам, в частности шифр “Квадрат” (Square) [12]. Но в процессе его разработки авторы придумали новую атаку и, чтобы обезопасить шифр от нее, вынуждены были увеличить рекомендуемое число раундов. Позднее под разными названиями аналогичные идеи появились в нескольких статьях, а обобщены были под названием *интегральный криптоанализ* [13].

При использовании интегрального криптоанализа исследуется то, как меняются свойства специальных мультимножеств при прохождении через раунды. Например, если мы установили, что множество, где каждый элемент встречается четное число раз, шифруется во множество с суммой, равной нулю после $R-1$ раундов, то по аналогичной схеме из предыдущего раздела можно найти подключ последнего раунда.

Интегральный криптоанализ ориентирован в основном на шифры, представляющие подстановочно-перестановочную сеть, и нечасто применяется для сети Фейстеля. Это атака по выбранному открытому тексту, так как здесь интерес представляют множества блоков с заданным свойством.

Помимо трех наиболее распространенных типов криптоанализа: линейного, дифференциального и интегрального — существуют другие подходы к вскрытию шифров, но они пока не дали таких серьезных результатов, как описанные методы, поэтому приведем лишь их основные идеи. Все алгебраические атаки имеют общую цель — представить шифровальный алгоритм в виде системы уравнений, где подключи являются неизвестными, и разрешить. Достаточно многообещающим в последние годы становится подход, основанный на том, что при физической реализации различные операции, участвующие в шифровании, потребляют разную мощность или выполняются разное время. Устройства для подобных измерений уже производятся и имеют вполне приемлемую цену — от нескольких сотен до нескольких тысяч долларов. Эти измерения часто в комбинации с вышеупомянутыми теоретическими атаками позволяют эффективно отыскивать подключи. Еще один вид атак называется *интерполяционным криптоанализом* [14], где на основе уменьшенного количества раундов шифра с помощью интерполяционных многочленов (например, Лагранжа) строится приближение на большее число раундов или на весь шифр.

4. Конкурсы блочных шифров

Из-за востребованности шифрования информации и достаточно большого числа предложенных шифров были организованы несколько крупных конкурсов, направленных на стандартизацию или исследование с дальнейшей рекомендацией к применению блочных шифров. Наиболее значительным оказался проект AES [5], ориентированный на поиск преемника DES как стандарта шифрования в США, а по сути и по всему миру. Требования к кандидатам были достаточно простыми: 128-битный блок, 128-, 192- и 256-битный секретный ключ, шифр должен по надежности не уступать 3-DES, но быть существенно быстрее него. На конкурс было подано порядка 20 заявок, из которых 15 шифров полностью соответствовали требованиям конкурса. После первого этапа были отобраны пять шифров, которые показали эффективную работу на различных платформах и хорошую надежность. Они исследовались еще более тщательно, но каких бы то ни было серьезных криптоаналитических результатов достигнуто не было и все шифры признались как “не имеющие недостатков”. Однако по условиям конкурса выбрать нужно было только один шифр, и в результате голосования, в котором участвовали спе-

циалисты, исследовавшие шифры, победил бельгийский Rijdael [6] во многом благодаря прозрачностии и элегантности дизайна.

После этого конкурса было проведено еще два — NESSIE [15] (в Европе) и CRYPTREC [16] (в Японии), где исследовались не только блочные шифры, но и другие алгоритмы.

Список литературы

[1] Shannon C. Communication Theory of secrecy systems // Bell System Technical J. 1949. Vol. 28. P. 656–715.

[2] Diffie W., Hellman M. New directions in cryptography // IEEE Trans. on Information Theory. 1976. Vol. 22 (6). P. 644–654.

[3] Feistel H. Cryptography and computer privacy // Sci. American. 1973. Vol. 228, N 5. P. 15–23.

[4] National Bureau of Standards. Data encryption standard // Federal Information Proc. Standard (FIPS). 1977. Vol. 81.

[5] Advanced Encryption Algorithm (AES) Development Effort // 1997–2000. `http://csrc.nist.gov/encryption/aes`

[6] Daemen J., Rijmen V. The Rijndael block cipher // AES Submission. 1999. `http://csrc.nist.gov/encryption/aes/rijndael/Rijndael.pdf`

[7] Biham E., Shamir A. Differential cryptanalysis of the full 16-round DES // Proc. of Crypto'92, LNCS. B.; N.Y.: Springer-Verl., 1993. Vol. 740. P. 487–496.

[8] Matsui M. Linear cryptanalysis method for DES cipher // Proc. of Eurocrypt'93, LNCS. B.; N.Y.: Springer-Verl., 1994. Vol. 765. P. 205–218.

[9] Knudsen L. Truncated and higher order differentials // Proc. of Fast Software Encryption'94, LNCS. B.; N.Y.: Springer-Verl., 1995. Vol. 1008. P. 196–211.

[10] Wagner D. The boomerang attack // Proc. of Fast Software Encryption'99, LNCS. B.; N.Y.: Springer-Verl., 1999. Vol. 1636. P. 156–170.

[11] Biham E., Biryukov A., Shamir A. Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials // Proc. of Eurocrypt'99, LNCS. B.; N.Y.: Springer-Verl., 1999. Vol. 1592. P. 12–23.

[12] Daemen J., Knudsen L., Rijmen V. The block cipher square // Proc. of Fast Software Encryption'97, LNCS. B.; N.Y.: Springer-Verl., 1997. Vol. 1267. P. 149–165.

[13] Knudsen L., Wagner D. Integral cryptanalysis // Proc. of Fast Software Encryption'02, LNCS. B.; N.Y.: Springer-Verl., 2002. Vol. 2365. P. 629–632.

[14] Jakobsen T., Knudsen L. The interpolation attack on block ciphers // Proc. of Fast Software Encryption'97, LNCS. B.; N.Y.: Springer-Verl., 1997. Vol. 1267. P. 28–40.

[15] New European Schemes for Signatures, Integrity, and Encryption // Deliverables of the NESSIE Project. 2003. `http://www.cosic.esat.kuleuven.ac.be/nessie`

[16] CRYPTREC Project // 2000–2002. `http://www.ipa.go.jp/security/enc/CRYPTREC`

Поступила в редакцию 4 июня 2007 г.