ИНСТИТУТ ИНТЕЛЛЕКТУАЛЬНЫХ КИБЕРНЕТИЧЕСКИХ СИСТЕМ

Кафедра
«Криптография и безопасность компьютерных систем»

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
к научно-исследовательской работе
«Обобщение интегральной атаки на основе инвариантных подпространств на семейство Khazad-подобных блочных шифров»

Исполнитель:
Студент гр. Б20-525 _____ Лебедев А.А.
подпись, дата

Научный руководитель: _____ Пудовкина М.А.
подпись, дата

Зам. зав. каф. № 41: _____ Пудовкина М.А.
подпись, дата

Москва – 2022

РЕФЕРАТ

Отчёт 22 с., 6 рис., 3 табл., 16 источников, 1 прил.

АЛГОРИТМ ШИФРОВАНИЯ, БЛОЧНЫЙ ШИФР, ИНВАРИАНТНЫЕ ПОДПРОСТРАНСТВА, ИНТЕГРАЛЬНАЯ АТАКА, РАУНДОВЫЙ КЛЮЧ, СЛАБЫЕ КЛЮЧИ, СХЕМА ФЕЙСТЕЛЯ

Объектом исследования является семейство Khazad-подобных блочных шифров, полученных из усечённого до 5 раундов шифра Khazad путём обобщения структуры линейного слоя, табличной нелинейной замены, размеров ключа, блока и подблока, а также алгоритма распространения ключа.

Целью работы является изучение структурных слабостей блочного шифра Khazad с уменьшенным числом раундов, а также границ применимости и вычислительной сложности некоторых существующих атак на данный шифр. В данной работе рассматривается атака на основе инвариантных подпространств и её применимость в случае возможного обобщения конфигурации и отдельных алгоритмов шифрсистемы.

Работа выполняется на основании данного задания по исследованию криптостойкости блочной шифрсистемы Khazad и возможности улучшения временных характеристик некоторых существующих атак на данный шифр.

СОДЕРЖАНИЕ

ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ

В настоящем отчете применяются следующие термины с соответствующими определениями, обозначениями и сокращениями:

$\overline{0,N}$ — множество целых чисел от 0 до N включительно.

|| — конкатенация векторов или строк конечной длины.

P — блок открытого текста.

$V_R(2^m)$ — множество векторов длины R над полем порядка 2^m.

$f \cdot g$ — произведение функций f и g, такое, что $(f \cdot g)(x) = g(f(x))$.

$W \oplus c$ — множество из сумм элементов пространства W с константой c

$(x)^f$ — применение функции f к x, т.е. значение $f(x)$.

$dim\, W$ — размерность пространства W.

$Binom(n, p)$ — биномиальное распределение с параметрами n и p.

MDS — maximal distance separable.

MDS-матрица — максимально рассеивающая матрица.

s-бокс — нелинейная подстановка, задаваемая таблицей замен.

Инвариантное подпространство — подпространство векторов, замкнутое относительно данного преобразования.

ЭВМ — электронная вычислительная машина.

ВВЕДЕНИЕ

С развитием вычислительной техники в криптографии стал всё шире применяться теоретико-информационный подход, и на смену традиционным шифрам пришли вычислительно стойкие шифры, рассчитанные на использование в ЭВМ и обеспечивающее стойкость к атакам с применением высокопроизводительных вычислительных устройств [1]. Так, наиболее широкое распространение на практике получили алгоритмы блочного симметричного шифрования, так как они допускают различные варианты конфигурации шифрсистемы и позволяют зашифровать текст произвольной длины на фиксированном ключе без существенной потери в криптостойкости [2].

На протяжении более 50 лет научное сообщество в области криптографии, а также государственные институты по стандартизации проводили конкурсы алгоритмов шифрования для установления новых стандартов и исследования различных шифрсистем. Одним из таких конкурсов стал проведённый в 2000 г. проект NESSIE (New European Schemes for Signature, Integrity, and Encryption), принявший 17 алгоритмов блочного симметричного шифрования [3].

Настоящая работа посвящена исследованию криптостойкости одного из финалистов данного конкурса — блочной шифрсистемы Khazad. Она была разработана криптографами Винсентом Риджменом и Пауло Баррето специально для участия в NESSIE. Отличительная особенность шифра заключается в инволютивности используемых преобразований, что в случае аппаратной реализации даёт возможность использовать одни и те же схемы для процедур зашифрования и расшифрования [4]. Данные процедуры у Khazad отличаются лишь в обратном порядке раундовых ключей в ключевом расписании и потому их реализацию можно выполнить с исключительной эффективностью использования пространства на плате и расходуемых логических элементов. Khazad был разработан в соответствии со стратегией

Wide Trail [5], согласно которой раундовая функция шифра состоит из обратимых преобразований, каждое из которых отвечает за отдельные криптографические свойства функции.

Линейный рассеивающий слой гарантирует зависимость каждого отдельного выходного бита одновременно от всех входных бит по прошествии нескольких раундов [6]. Нелинейный слой посредством удаления функции от линейной обеспечивает соответственно достаточную стойкость к линейному криптоанализу. Подмешивание ключа происходит посредством побитового сложения с соответствующим раундовым ключом по модулю 2.

На данный момент известны атаки на 3, 4 и 5 раундов шифра, среди которых:

- интегральная атака [4] на 3 раунда (2^{16} операций применений s-бокса, материал в 2^9 блоков открытых и шифрованных текстов), на 4 раунда (2^{80} и 2^9 соответственно);
- атака на основе запретных разностей [7] на 3 раунда с трудоёмкостью 2^{64} операций расшифрования за один раунд при объёме материала 2^{13} блоков открытого и шифртекстов;
- слайд-атака [8], существенно использующая инволютивность преобразований, позволяет вскрыть 5 раундов системы при использовании ключа из класса слабых ключей мощностью 2^{64} с трудоёмкостью 2^{43} операций обращения в память и материале 2^{38} пар блоков открытого и шифртекстов;
- интегральная атака на основе инвариантных подпространств [9], позволяющая при использовании ключа из одного из 7 классов слабых ключей мощностью 2^{64} или 7 классов по 2^{32} каждый восстановить раундовый ключ с трудоёмкостью 2^{43} при объёме материала 2^{32}, или 2^{27} и 2^{16} соответственно;
- некоторые другие виды атак, имеющие примерно схожую с предыдущими трудоёмкость.

В настоящем отчёте рассмотрена интегральная атака на основе инвариантных подпространств и соответствующие классы слабых ключей. Сделаны некоторые обобщения относительно алгоритма развёртывания ключа, используемых преобразований, а также размеров ключа и блока. Рассмотрена возможность применения атаки на обобщённый вариант шифрсистемы.

1 Описание блочного шифра Khazad

Шифр Khazad использует инволютивные преобразования, благодаря чему процедуры зашифрования и расшифрования у Khazad отличаются лишь в обратном порядке раундовых ключей в ключевом расписании [4]. Несмотря на преимущества подобного подхода, это может привести к потенциальным слабостям шифрсистемы.

1.1 Используемые алгебраические структуры

Данный шифр оперирует байтами, представленными элементами конечного поля $GF(2^8)$. Это поле представляется как $GF(2)[x]/p(x)$, где $p(x)=x^8+x^4+x^3+x^2+1$ — первый неприводимый многочлен из списка [10]. Многочлен $p(x)$ был выбран так, чтобы $g(x)=x$ был генератором мультипликативной группы $GF(2^8)\setminus\{0\}$.

Элемент $u=u_7x^7+u_6x^6+...+u_1x+u_0$, где $u_i \in GF(2), i=\overline{0,7}$, обозначим шестнадцатиричным числом $u=u_7\cdot 2^7+u_6\cdot 2^6+...+u_1\cdot 2+u_0$, взятым в одинарные кавычки (например, $u=x^2+1$ обозначается как '5', а многочлену $p(x)$ будет соответствовать '11d').

1.2 Используемые преобразования

Khazad — блочная шифрсистема со 128-битным ключом и длиной блока 64 бита, представляемым в виде вектора $\alpha=(a_0,...,a_7) \in V_8(2^8)$.

Шифр использует три основных преобразования:

- линейное h: умножение на максимально рассеивающую матрицу Коши [11] $H=\|h_{i,j}\| \in GL_8(2^8)$ (рис. 1);
- нелинейное $\tilde{s}$: применение табличной замены (S-бокса) к координатам вектора (приведена в Приложении А);
- наложение ключа ν_k: побитовое сложение по модулю 2 с соответствующим байтом ключа k.

$$H = \begin{bmatrix} \text{'01'} & \text{'03'} & \text{'04'} & \text{'05'} & \text{'06'} & \text{'08'} & \text{'0b'} & \text{'07'} \\ \text{'03'} & \text{'01'} & \text{'05'} & \text{'04'} & \text{'08'} & \text{'06'} & \text{'07'} & \text{'0b'} \\ \text{'04'} & \text{'05'} & \text{'01'} & \text{'03'} & \text{'0b'} & \text{'07'} & \text{'06'} & \text{'08'} \\ \text{'05'} & \text{'04'} & \text{'03'} & \text{'01'} & \text{'07'} & \text{'0b'} & \text{'08'} & \text{'06'} \\ \text{'06'} & \text{'08'} & \text{'0b'} & \text{'07'} & \text{'01'} & \text{'03'} & \text{'04'} & \text{'05'} \\ \text{'08'} & \text{'06'} & \text{'07'} & \text{'0b'} & \text{'03'} & \text{'01'} & \text{'05'} & \text{'04'} \\ \text{'0b'} & \text{'07'} & \text{'06'} & \text{'08'} & \text{'04'} & \text{'05'} & \text{'01'} & \text{'03'} \\ \text{'07'} & \text{'0b'} & \text{'08'} & \text{'06'} & \text{'05'} & \text{'04'} & \text{'03'} & \text{'01'} \end{bmatrix}$$

Рисунок 1 — используемая в линейном преобразовании матрица Коши

В спецификации шифра [4] показано, что матрица H приводима, симметрична и унитарна, следовательно, инволютивна.

Используемый s-бокс был подобран так, чтобы также обладать инволютивным свойством, т.е. $s[s[x]]=x \forall\, x \in GF(2^8)$. Так, нелинейный слой также является инволюцией.

Подмешивание ключа, реализованное побитовой операцией сложения по модулю 2, очевидно, тоже инволютивно.

1.3 Раундовая функция и ключевое расписание

Шифр использует набор раундовых констант: для i-го раунда вектор $c^i \in GF(2^8)^8$ определён как

$$c^i_j = s[8i+j], 0 \le j \le R,$$

где R — размер блока шифртекста, равный 8 в оригинальной шифрсистеме.

Раундовая функция r-го раунда является отображением $\rho[k]: GF(2^8)^8 \to GF(2^8)^8$, параметризованным раундовым ключом $k \in V_8(2^8)$, и определяется как

$$\rho[k] \equiv \tilde{s} \cdot h \cdot v_k.$$

Тогда функция зашифрования за r раундов может быть записана как

$$g_{k_0,k_1,\ldots,k_r} = v_0 \tilde{s} h v_{k_1} \tilde{s} h v_{k_2} \ldots \tilde{s} h v_{k_{r-1}} \tilde{s} v_{k_r}. \quad (1)$$

Ключ шифрования представляется в виде двух векторов $k_{-1}, k_{-2} \in V_8(2^8)$. Алгоритм развёртывания ключа (составления ключевого расписания) реализуется как зашифрование векторов k_{-1}, k_{-2} на сети Фейстеля [12] с функцией усложнения

$$f_{c_i} = \tilde{s} h v_{c_i}. \quad (2)$$

Таким образом, раундовые ключи вычисляются следующим образом:

$$k_i = \left(k_{i-1}\right)^{f_{c_i}} \oplus k_{i-2}. \quad (3)$$

2 Принципы интегральной атаки на основе инвариантных подпространств

В данном разделе описана идея интегральной атаки на основе инвариантных подпространств [9], рассмотрены используемые ей слабости шифра, приведена оценка эффективности данной атаки для разных размеров подпространств.

2.1 Свойства раундовых преобразований шифра

Опишем некоторые свойства, связанные с используемыми в Khazad преобразованиями. Эти свойства связаны в первую очередь с выбором матрицы H. Так как в оригинальном шифре матрица H максимально рассеивающая и приводимая, в сочетании с выбором s-боксов с высокими криптографическими параметрами это даёт устойчивость даже 5 раундов шифра к линейному и дифференциальному криптоанализу.

Однако потенциальной слабостью алгоритма является наличие у преобразований инвариантных подпространств — таких пространств векторов, которые замкнуты относительно данного преобразования. Так, в работе [9] было обнаружено 7 инвариантных относительно $\tilde{s}$ и h подпространств $V_8(2^8)$ мощностью по 2^{32} каждое, 7 подпространств по 2^{16} векторов и 1 подпространство мощностью 2^8 (рис. 2). Так как данные подпространства инвариантны относительно $\tilde{s}$ и h, в ходе процедуры зашифрования только преобразование ν_k может вывести вектор из такого подпространства. При этом ν_k оставляет вектор в том же подпространстве, если k также находится в нём.

2.2 Описание интегральной атаки на 5 раундов

Пусть W — одно из перечисленных ранее подпространств. Поскольку в алгоритме развёртывания ключа используется биективная раундовая функция, существует ровно $|W|^2$ пар векторов (k_{-1}, k_{-2}) таких, что $k_1, k_2 \in W$.

Пусть при зашифровании используются k_{-1}, k_{-2}, удовлетворяющие этому свойству. Так как $k_2 = (k_1)^{\tilde{s}hv_{c_0}} \oplus k_0 = (k_1)^{\tilde{s}h} \oplus (k_0 \oplus c_2) \in W$, получаем $k_0 \in W \oplus c_2$;

Аналогично $k_3 = (k_2)^{\tilde{s}hv_{c_3}} \oplus k_1 \in W$, тогда $k_3 \in W \oplus c_3$.

Тогда можно заметить, что после трёх раундов шифрования все блоки открытого текста из смежного класса $W \oplus c_2$ переходят в класс $W \oplus c_3$, так что

$$(W \oplus c_2)^{v_{k_0}\tilde{s}hv_{k_1}\tilde{s}hv_{k_2}\tilde{s}hv_{k_3}} = W \oplus c_3.$$

Класс $W \oplus c_3$ не является инвариантным, поэтому преобразования $\tilde{s}$ и h не обязательно сохранят его.

Таким образом, открывается возможность применить к 5 раундам шифра интегральную атаку, используя проверку данного равенства.

Трудоёмкость алгоритма можно оценить в $8 \cdot 2^8 |W|$ операций $\tilde{s}$ при объеме материала $|W|$ блоков текста. Метод оказывается эффективным, когда $8 \cdot 2^8 |W| < |W|^2$, т.е. при $|W| > 2^{11}$. Отсюда следует, в частности эффективность метода для подпространств 1 – 14 и неэффективность для подпространства 15 рисунка 2 по сравнению с полным перебором по соответствующему классу слабых ключей.

Так, для подпространств мощностью 2^{32} трудоёмкость составит 2^{43} операций при объеме в 2^{32} блоков текста;

Для подпространств мощностью 2^{16} трудоёмкость составит 2^{27} операций при объеме в 2^{16} блоков текста.

3 Обобщения конфигурации шифрсистемы

В данном разделе предложены обобщения параметров шифра Khazad, в совокупности образующие семейство Khazad-подобных шифров. Рассмотрены факторы, влияющие на образование инвариантных подпространств.

3.1 Размер ключа, блока и подблока

В алгоритме шифрования Khazad размер блока равен 64 битам, что соответствует $N=8$ подблокам (байтам) по $M=8$ бит. Соответственно, размерность матрицы H составляет $N \times N$, т.е. 8×8, а все операции происходят над полем $GF(2^M)$. Пусть ключ состоит из r блоков (в Khazad $r=2$), тогда размер ключа составит rMN бит. Далее обобщим идею атаки для произвольных r, M, N и произвольной MDS-матрицы H.

3.2 Подпространства, инвариантные относительно s- и h-преобразований

Центральным понятием в рассматриваемой атаке являются подпространства, инвариантные одновременно к применению s-бокса и к умножению на произвольную MDS-матрицу.

Отметим, что, поскольку s-бокс (суть произвольная подстановка $s: GF(2^M) \to GF(2^M)$) применяется к координатам вектора по отдельности, любые подпространства, определяемые непосредственной зависимостью одних координат от других, будут инвариантны относительно s. На основе этих рассуждений формулируем следующее утверждение.

Утверждение 1. *Любое подпространство, образованное непосредственной зависимостью каких-либо координат вектора, инвариантно относительно применения произвольного s-бокса.*

Доказательство. Пусть W_0 — подпространство $V_N(2^M)$, в котором для определенности первые $p(1 \le p \le N)$ координат независимы, а остальные $(N-p)$ заполняются циклически значениями первых p координат, т.е.

$$W_0 = \left\{ (x_1, \ldots, x_p, x_1, \ldots, x_p, \ldots, x_1, \ldots, x_{1+(N-1 \bmod p)}) \mid x_1, \ldots, x_p \in V_N(2^M) \right\}. \quad (4)$$

Применим подстановку s к координатам W_0:

$$W_s = \left\{ \left(x_1^s, \ldots, x_p^s, \ldots, x_1^s, \ldots, x_{1+(N-1 \bmod p)}^s \right) \mid x_1, \ldots, x_p \in V_N\left(2^M\right) \right\}. \quad (5)$$

Очевидно, после применения s к координатам с одинаковыми значениями получаются также одинаковые между собой значения. Тогда структура подпространства сохраняется, и в силу биективности s оказывается, что два множества совпадают с точностью до расстановки векторов. Получаем, что $W_0 = W_s$.

Замкнутость W_0 относительно операций сложения и умножения на элемент поля $GF\left(2^M\right)$, не равный нулю, проверяется аналогичным образом. Для нулевого элемента замкнутость проверяется непосредственно.

Для любых других позиций независимых координат и количественном соотношении тех или иных координат в векторе рассуждения полностью аналогичны. ∎

Природа подпространств, имеющихся у линейных преобразований, несколько сложнее и по большей части зависит от выбора матрицы. Однако в алгоритмах шифрования часто используются рекурсивные матрицы с идентичными или «отзеркаленными» блоками [14], и благодаря подобной структуре обнаружение подпространств, образующихся непосредственной зависимостью координат, т.е. подобных имеющимся в применении s-бокса, существенно облегчается.

В настоящем отчёте не рассматривается природа возникновения инвариантных подпространств в линейном слое h. Будем полагать, что имеется некоторое подпространство $W \subset V_N\left(2^M\right)$, инвариантное относительно преобразований $\tilde{s}$ и h, независимо от его происхождения.

3.3 Алгоритм развёртывания ключа

В шифрсистеме Khazad алгоритм развёртывания ключа (составления ключевого расписания) реализуется как зашифрование начальных векторов k_{-1}, k_{-2} на сети Фейстеля с функцией усложнения $f_{c_i} = \tilde{s} h v_{c_i}$.

Раундовые ключи вычисляются рекуррентно по формуле (3). Обобщим данную процедуру: пусть алгоритм развёртывания ключа работает на обобщённой сети Фейстеля [15] (рис. 4) с длиной регистра r, а ключ так же имеет длину r блоков.

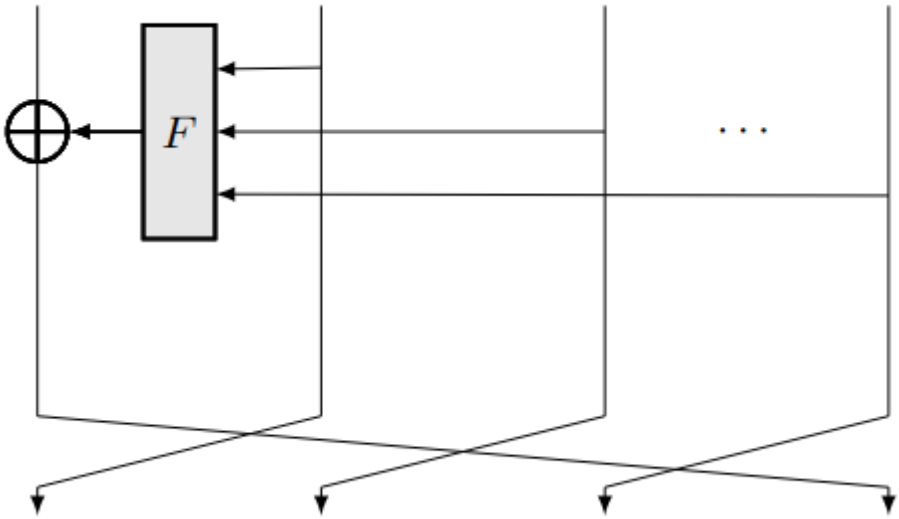

Рисунок 4 — пример обобщённой сети Фейстеля с общей раундовой функцией

Тогда процедуру вычисления i-го раундового ключа k_i можно описать уравнением

$$k_i = f_i\left(k_{i-r+1}, \ldots, k_{i-1}\right) \oplus k_{i-r}, \tag{6}$$

где f_i — i-я раундовая функция.

Функцию усложнения f_i, использующую раундовые преобразования процедуры зашифрования $\tilde{s}$, h и v_{c_i}, обобщим на случай нескольких переменных:

$$f_i\left(k_1, \ldots, k_n\right) = \oplus_{i=1}^{n}\left(k_i\right)^{\tilde{s}h} \oplus c_i. \tag{7}$$

4 Применимость интегральной атаки к обобщённой модели шифра

Исследуем применимость интегральной атаки к обобщённой модели шифра, принимая во внимание её отличия от оригинального шифра и возможные дополнительные поправки.

4.1 Обобщение алгоритма восстановления ключа

Пусть алгоритм составления ключевого расписания задаётся (6). В силу биективности преобразований $\tilde{s}$ и ν_k существуют $|W|^r$ значений для $(k_{-1},\dots,k_{-r})$ таких, что $k_1,\dots,k_r \in W$. Пусть $k_{-1},\dots,k_{-r}$ удовлетворяют этому свойству. Поскольку $k_r = \oplus_{i=1}^{r-1}(k_i)^{\tilde{s}h} \oplus c_r \oplus k_0 \in W$, имеем $W \oplus (c_r \oplus k_0) = W$, а значит $k_0 \in (W \oplus c_r)$. Далее применим рассуждения из раздела 2.2 к формуле распространения ключа (6) с функцией усложнения (7).

Проследим, в какие подпространства попадают блоки открытого текста $P \in (W \oplus c_r)$ в процессе $(r+1)$ раундов зашифрования.

- после первичного отбеливания ν_{k_0}: $(P)^{\nu_{k_0}} = P \oplus k_0 \in W \oplus c_r \oplus W \oplus c_r = W$;
- преобразование $\tilde{s}h$ сохраняет результат в W;
- сложение со следующим ключом ν_{k_1} также сохраняет W, т.к. $k_1 \in W$;
- последующие преобразования до $\nu_{k_{r+1}}$ аналогичным образом сохраняют W;
- сложение с ключом $\nu_{k_{r+1}}$, где $k_{r+1} \in W \oplus c_{r+1}$, переведёт вектор в $W \oplus c_{r+1}$.

Рассмотрим множество $A = (W \oplus c_{r+1})^{\tilde{s}h}$. Представим его в виде $\{a_i = (a_{i,0},\dots,a_{i,7}) | i = \overline{0, W-1}\}$. Тогда можно проверить равенство из [9], однако теперь требование о том, чтобы каждый элемент из $GF(2^M)$ встречался в любой фиксированной координате векторов из W чётное число раз, в общем случае может быть не выполнено. Поэтому сформулируем следующее обобщённое утверждение.

Утверждение 2. *Если каждый элемент из $GF(2^M)$ встречается в любой фиксированной координате векторов из W чётное количество раз, то для любого $j \in \overline{0, M-1}$ справедливо равенство*

$$\sum_{i=0}^{2^u-1} a_{i,j}=0, \tag{8}$$

где $u=dim\,W$.

Схема доказательства данного утверждения остаётся идентичной доказательству в [9].

Замечание. Для корректности утверждения необходимо также, чтобы матрица H была максимально рассеивающей [11].

Таким образом, появляется возможность применить тот же алгоритм, что на рисунке 3, с поправкой на числовые обобщения.

Пусть $B=\left\{\alpha^{g[k_0,\ldots,k_{r+3}]} \mid \alpha \in W \oplus c_r\right\}$, $\beta_i=(b_{i,0},\ldots,b_{i,M-1}), i \in \overline{0,|W|-1}$, $u=dim\,W$. Следующий алгоритм (рис. 5) формирует множество вариантов ключа $k_{r+3} \in V_N(2^M)$.

```
KeySet ← (Ø, Ø, ..., Ø)_N
for i=0 to N do
    for k_{r+3,i}=0 to 2^M−1 do
        IntSum ← Σ_{i=0}^{2^u−1} (b_{j,i} ⊕ k_{r+3,i})^s
        if IntSum=0 then
            KeySet_i ← KeySet_i ∪ {k_{r+3,i}}
        end
    end
end
```

Рисунок 5 — модифицированный алгоритм восстановления (r+3)-го ключа

Для обобщённого алгоритма можно показать, что в каждом из N множеств *KeySet* получается в среднем не более 2 вариантов ключа.

Действительно, для случайной подстановки вероятность равенства нулю *IntSum* составит $p=2^{-M}$, тогда математическое ожидание 1 плюс суммы $n=2^M-1$ бернуллиевских величин (суть $Binom(n,p)$), оценивается как

$$E\left|KeySet_i\right|=1+np=1+2^{-M}\left(2^M-1\right)=2-2^{-M}\le 2, \tag{9}$$

и тогда остаётся лишь подобрать ключ из не более чем 2^N вариантов.

Таким образом, интегральная атака на основе подпространства W позволяет получить раундовый ключ для $(r+3)$ раундов шифра.

4.2 Оценка трудоёмкости алгоритма

По виду алгоритма нетрудно оценить трудоёмкость интегральной атаки. Для класса слабых ключей мощностью $|W|^r$ объём необходимого материала составляет $|W|$ блоков текста. Перемножая размерности вложенных циклов в алгоритме, для временной трудоёмкости получаем $N \cdot 2^M |W|$ операций применения s-бокса.

Данная атака оказывается эффективнее полного перебора по $|W|^r$ вариантам ключа, если $N \cdot 2^M |W| < |W|^r$, т.е.

$$|W| > \sqrt[r-1]{N} \cdot 2^{\frac{M}{r-1}}. \qquad (10)$$

Стоит отметить, что с увеличением мощности подпространства W пропорционально растёт объём необходимых пар открытого и шифрованного текстов, однако увеличивается выигрыш во временной сложности по сравнению с атакой методом полного перебора по соответствующему классу слабых ключей.

ЗАКЛЮЧЕНИЕ

В настоящем отчёте была проанализирована интегральная атака на основе инвариантных подпространств на 5 раундов шифра Khazad, а также рассмотрена её применимость к обобщённой модели шифра — семейству Khazad-подобных шифров с произвольным размером ключа, блока и подблока, а также с модифицированным алгоритмом составления ключевого расписания.

Было сформулировано и рассмотрено обобщение атаки на данное семейство шифров для инвариантного подпространства W и получены следующие результаты:

- при размере ключа в r блоков атака позволяет восстановить $(r+3)$-й раундовый ключ для некоторого класса слабых ключей $K \subset \left(V_N\left(2^M\right)\right)^r$;
- необходимый объём материала составляет $|W|$ пар открытого и шифрованного текстов;
- временная сложность атаки составляет $N \cdot 2^M |W|$ операций применения s-бокса;
- надёжность метода равна 1;
- число слабых ключей равно $|W|^r$.

На подпространство W накладывается дополнительное требование: каждый элемент поля $GF\left(2^M\right)$ должен встречаться в любой фиксированной координате векторов из W чётное число раз.

При мощности инвариантного подпространства, удовлетворяющей неравенству (10), данная атака оказывается эффективнее полного перебора по соответствующему классу слабых ключей.

Дальнейшие исследования ранее рассмотренной атаки могут быть направлены на улучшение её временных характеристик, изучение границ её применимости при дальнейшем обобщении конфигурации шифра, а также изучение методов нахождения новых инвариантных подпространств и идентификации слабых ключей.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. От манускриптов до шифровальных машин: история криптографии [Электронный ресурс] // naked-science.ru: [сайт]. [2019]. URL: https://naked-science.ru/article/sci/ot-manuskriptov-do-shifrovalnyh (дата обращения: 13.12.2022).

2. Block Cipher [Электронный ресурс] // wikipedia.org: [сайт]. [2022]. URL: https://en.wikipedia.org/wiki/Block_cipher (дата обращения 13.12.2022)

3. New European Schemes for Signatures, Integrity, and Encryption [Электронный ресурс] // nessie.eu.org: [сайт]. [2003]. URL: http://www.nessie.eu.org/index.html (дата обращения: 13.12.2022).

4. Barreto P., Rijmen V. The Khazad Legacy-Level Block Cipher. In First Open NESSIE Workshop // KU Leuven, 2000. URL: https://www.cosic.esat.kuleuven.be/nessie/workshop/submissions/khazad.zip (дата обращения: 13.12.2022).

5. J. Daemen. Cipher and hash function design strategies based on linear and differential cryptanalysis // KU Leuven, 1995. С. 115-122. URL: https://cs.ru.nl/~joan/papers/JDA_Thesis_1995.pdf (дата обращения: 13.12.2022).

6. Chand Gupta, K., Ghosh Ray, I. On constructions of involutory MDS matrices. In: Youssef, A., Nitaj, A., Hassanien, A.E. (eds.) // AFRICACRYPT 2013. LNCS, vol. 7918, pp. 43–60. Springer, Heidelberg (2013).

7. NESSIE Security Report D20, version 2-0 // NESSIE project, 2003. URL: https://www.cosic.esat.kuleuven.be/nessie/deliverables/D20-v2.pdf (дата обращения: 13.12.2022).

8. Biryukov A. Analysis of Involutional Ciphers: Khazad and Anubis. Fast Software Encryption // LNCS, 2003. Springer (2003).

9. Буров Д.А., Погорелов Б.А. Атака на пять раундов шифрсистемы Khazad // Математические вопросы криптографии, 2016, Т. 7, № 2, С. 35–46.

10. R. Lidl, H. Hiederreiter. Introduction to finite fields and their applications // Cambridge University Press, 1986.

11. Sim, S.M., Khoo, K., Oggier, F., Peyrin, T. (2015). Lightweight MDS Involution Matrices. In: Leander, G. (eds) // Fast Software Encryption. FSE 2015. Lecture Notes in Computer Science, vol 9054. Springer, Berlin, Heidelberg.

12. Menezes, Alfred J.; Oorschot, Paul C. van; Vanstone, Scott A. Handbook of Applied Cryptography (Fifth ed.). p. 251. ISBN 978-0849385230.

13. Binomial distribution [Электронный ресурс] // wikipedia.org: [сайт]. [2022]. URL: https://en.wikipedia.org/wiki/Binomial_distribution (дата обращения: 18.12.2022).

14. Albrecht, M.R., Driessen, B., Kavun, E.B., Leander, G., Paar, C., Yalçın, T. (2014). Block Ciphers – Focus on the Linear Layer (feat. PRIDE) In: Garay, J.A., Gennaro, R. (eds) // Advances in Cryptology – CRYPTO 2014. CRYPTO 2014. Lecture Notes in Computer Science, vol 8616. Springer, Berlin, Heidelberg.

15. Hoang, V.T., Rogaway, P. On Generalized Feistel Networks. In: Rabin, T. (eds) // Advances in Cryptology – CRYPTO 2010. CRYPTO 2010. Lecture Notes in Computer Science, vol 6223. Springer, Berlin, Heidelberg.

16. Панасенко С. П. Алгоритмы шифрования. Специальный справочник. // СПб.: БХВ-Петербург, 2009. С. 282–287. ISBN 978-5-9775-0319-8.

Приложение А

Таблица замен используемого в шифре S-бокса

Пользуясь возможностью незначительного изменения алгоритма в течение первого раунда NESSIE, авторы Khazad внесли изменения в свой алгоритм. В новом варианте спецификации исходный алгоритм Khazad назван Khazad-0, а название Khazad присвоено модифицированному алгоритму [16].

В оригинальном шифре KHAZAD-0 табличная замена представлялась классическим S-боксом и описывалась следующей таблицей замен (табл. А.1).

Таблица А.1 — исходный S-бокс шифра Khazad-0

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	A7	D3	E6	71	D0	AC	4D	79	3A	C9	91	FC	1E	47	54	BD
1	8C	A5	7A	FB	63	B8	DD	D4	E5	B3	C5	BE	A9	88	0C	A2
2	39	DF	29	DA	2B	A8	CB	4C	4B	22	AA	24	41	70	A6	F9
3	5A	E2	B0	36	7D	E4	33	FF	60	20	08	8B	5E	AB	7F	78
4	7C	2C	57	D2	DC	6D	7E	0D	53	94	C3	28	27	06	5F	AD
5	67	5C	55	48	0E	52	EA	42	5B	5D	30	58	51	59	3C	4E
6	38	8A	72	14	E7	C6	DE	50	8E	92	D1	77	93	45	9A	CE
7	2D	03	62	B6	B9	BF	96	6B	3F	07	12	AE	40	34	46	3E
8	DB	CF	EC	CC	C1	A1	C0	D6	1D	F4	61	3B	10	D8	68	A0
9	B1	0A	69	6C	49	FA	76	C4	9E	9B	6E	99	C2	B7	98	BC
A	8F	85	1F	B4	F8	11	2E	00	25	1C	2A	3D	05	4F	7B	B2
B	32	90	AF	19	A3	F7	73	9D	15	74	EE	CA	9F	0F	1B	75
C	86	84	9C	4A	97	1A	65	F6	ED	09	BB	26	83	EB	6F	81
D	04	6A	43	01	17	E1	87	F5	8D	E3	23	80	44	16	66	21
E	FE	D5	31	D9	35	18	02	64	F2	F1	56	CD	82	C8	BA	F0
F	EF	E9	E8	FD	89	D7	C7	B5	A4	2F	95	13	0B	F3	E0	37

В модифицированном варианте для уменьшения занимаемой s-боксом площади на микросхемах (в случае аппаратной реализации) используется комбинированный S-бокс из блоков P и Q, каждый из которых является маленьким блоком замены с 4 битами на входе и выходе (рис. А.1).

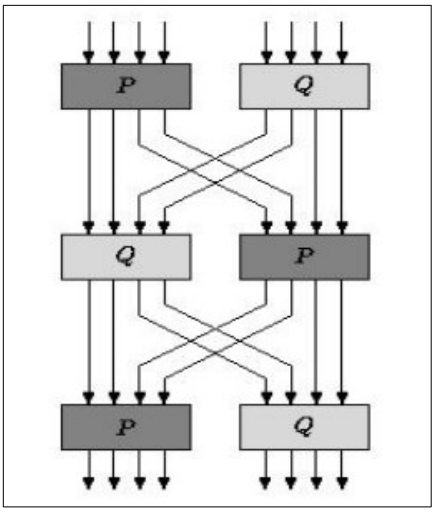

Рисунок А.1 — схема модифицированного S-бокса шифра

Блоки P и Q задаются следующими таблицами замен (табл. А.2).

Таблица А.2 — таблица замен блоков P и Q

x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
P(x)	3	F	E	0	5	4	B	C	D	A	9	6	7	8	2	1

x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
Q(x)	9	E	5	6	A	2	3	C	F	0	4	D	7	B	1	8

Данная схема эквивалентна применению следующего S-бокса (табл. А.3).

Таблица А.3 — результирующая таблица замен модифицированного S-бокса

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	BA	54	2F	74	53	D3	D2	4D	50	AC	8D	BF	70	52	9A	4C
1	EA	D5	97	D1	33	51	5B	A6	DE	48	A8	99	DB	32	B7	FC
2	E3	9E	91	9B	E2	BB	41	6E	A5	CB	6B	95	A1	F3	B1	02
3	CC	C4	1D	14	C3	63	DA	5D	5F	DC	7D	CD	7F	5A	6C	5C
4	F7	26	FF	ED	E8	9D	6F	8E	19	A0	F0	89	0F	07	AF	FB
5	08	15	0D	04	01	64	DF	76	79	DD	3D	16	3F	37	6D	38
6	B9	73	E9	35	55	71	7B	8C	72	88	F6	2A	3E	5E	27	46
7	0C	65	68	61	03	C1	57	D6	D9	58	D8	66	D7	3A	C8	3C
8	FA	96	A7	98	EC	B8	C7	AE	69	4B	AB	A9	67	0A	47	F2
9	B5	22	E5	EE	BE	2B	81	12	83	1B	0E	23	F5	45	21	CE
A	49	2C	F9	E6	B6	28	17	82	1A	8B	FE	8A	09	C9	87	4E
B	E1	2E	E4	E0	EB	90	A4	1E	85	60	00	25	F4	F1	94	0B
C	E7	75	EF	34	31	D4	D0	86	7E	AD	FD	29	30	3B	9F	F8
D	C6	13	06	05	C5	11	77	7C	7A	78	36	1C	39	59	18	56
E	B3	B0	24	20	B2	92	A3	C0	44	62	10	B4	84	43	93	C2
F	4A	BD	8F	2D	BC	9C	6A	40	CF	A2	80	4F	1F	CA	AA	42