Тим Бейн
Винсент Рэймен

ЛИНЕЙНЫЙ КРИПТОАНАЛИЗ

$$u_{1,i} + v_{2,i} + v_{3,i} + v_{4,i} = 4_{d1} - 2 \quad {}^{2}{}_{,i3},$$
$$u_{2,i} + v_{1,i} + v_{3,i} + v_{4,i} =$$
$$u_{3,i} + v_{1,i} + v_{2,i} +$$
$$u_{4,i} + v_{1,i} + v_{2,i} +$$

$$\begin{matrix}
4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -2 & 0 & 0 & 0 & 0 & 0 & -2 \\
0 & 0 & -2 & 0 & 0 & 2 & 0 & -2 \\
0 & -2 & 2 & 0 & 0 & 2 & 2 & 0 \\
0 & 0 & 0 & 0 & -2 & 2 & -2 & -2 \\
0 & 2 & 0 & 2 & -2 & 0 & 2 & 0 \\
0 & 0 & -2 & 2 & 2 & 2 & 0 & 0 \\
0 & -2 & -2 & 0 & -2 & 0 & 0 & 2
\end{matrix}$$

ДМК ПРЕСС
ИЗДАТЕЛЬСТВО

Тим Бейн, Винсент Рэймен

Линейный криптоанализ

Linear Cryptanalysis

Tim Beyne, Vincent Rijmen

Линейный криптоанализ

Тим Бейн, Винсент Рэймен

ДМК ПРЕСС
издательство
Москва, 2026

УДК 003.26
ББК 16.8
Б41

Тим Бейн, Винсент Рэймен
Б41 Линейный криптоанализ / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2026. – 180 с.: ил.

ISBN 978-5-93700-474-1

Данное руководство посвящено анализу безопасности (криптоанализу) фундаментальных блоков, на которых основаны криптографические приложения. Линейный криптоанализ рассматривается с математической точки зрения и сопровождается обзором наиболее влиятельных публикаций. Главы дополнены большим количеством примеров и упражнений, опирающихся на теорию и практику.

Предварительные знания теории криптографии не требуются. Издание будет полезно как начинающим читателям, изучающим криптографию, так и опытным экспертам, применяющим ее на практике.

УДК 003.26
ББК 16.8

ISBN (анг.) 978-1-00960-786-5
ISBN (рус.) 978-5-93700-474-1

Оглавление

Предисловие от издательства

Отзывы и пожелания

Мы всегда рады отзывам наших читателей. Расскажите нам, что вы думаете об этой книге – что понравилось или, может быть, не понравилось. Отзывы важны для нас, чтобы выпускать книги, которые будут для вас максимально полезны.

Вы можете написать отзыв на нашем сайте *www.dmkpress.com*, зайдя на страницу книги и оставив комментарий в разделе «Отзывы и рецензии». Также можно послать письмо главному редактору по адресу *dmkpress@gmail.com*; при этом укажите название книги в теме письма.

Если вы являетесь экспертом в какой-либо области и заинтересованы в написании новой книги, заполните форму на нашем сайте по адресу *http://dmkpress.com/authors/publish_book/* или напишите в издательство по адресу *dmkpress@gmail.com*.

Список опечаток

Хотя мы приняли все возможные меры для того, чтобы обеспечить высокое качество наших текстов, ошибки все равно случаются. Если вы найдете ошибку в одной из наших книг – возможно, ошибку в основном тексте или программном коде, – мы будем очень благодарны, если вы сообщите нам о ней. Сделав это, вы избавите других читателей от недопонимания и поможете нам улучшить последующие издания этой книги.

Если вы найдете какие-либо ошибки в коде, пожалуйста, сообщите о них главному редактору по адресу *dmkpress@gmail.com*, и мы исправим это в следующих тиражах.

Нарушение авторских прав

Пиратство в интернете по-прежнему остается насущной проблемой. Издательство «ДМК Пресс» очень серьезно относится к вопросам защиты авторских прав и лицензирования. Если вы столкнетесь в интернете с незаконной публикацией какой-либо из наших книг, пожалуйста, пришлите нам ссылку на интернет-ресурс, чтобы мы могли применить санкции.

Ссылку на подозрительные материалы можно прислать по адресу *dmkpress@gmail.com*.

Мы высоко ценим любую помощь по защите наших авторов, благодаря которой мы можем предоставлять вам качественные материалы.

Предисловие

Криптоанализ остается молодой и быстро развивающейся областью знаний. Поэтому лекторам и их ассистентам часто бывает трудно найти подходящие учебники для эффективного преподавания теории и практики студентам. Данной книгой мы надеемся закрыть этот пробел хотя бы в части линейного криптоанализа.

На наш взгляд, из всех методов криптоанализа шифров с симметричным ключом именно линейный криптоанализ лучше подходит для первокурсников. Интуитивно понятно, как строятся и выполняются линейные атаки. В то же время для научного описания линейного криптоанализа необходимы некоторые базовые, а иногда даже продвинутые сведения из линейной алгебры.

О чем эта книга

Вы можете рассчитывать, что, тщательно изучив эту книгу, получите глубокое понимание базовой теории линейного криптоанализа и познакомитесь с ее наиболее важными обобщениями (множественный и многомерный линейный криптоанализ, линейный криптоанализ с нулевой корреляцией и т. д.). Если вы также будете прилежно решать упражнения, то сможете применить полученные знания на практике. И тогда в нужное время у вас не возникнет проблем с пониманием современной литературы.

Тем не менее в книге такого размера невозможно не то что дать полный обзор всех работ на эту тему, но хотя бы перечислить их. Поэтому мы сознательно сделали упор на базовые криптоаналитические результаты, а многие важные, но имеющие лишь косвенное отношение к теме вопросы (например, связь с линейными кодами, булевы функции и т. д.) вынесли в упражнения. Большинство примеров и упражнений относятся к блочным шифрам, но мы ожидаем, что вы сможете применить линейный криптоанализ и к другим криптографическим примитивам, таким как потоковые шифры.

Стремясь избежать обвинения в неосновательном фаворитизме, мы осознанно ограничились короткими списками литературы. Они приведены в конце каждой главы. Отбор ссылок основан на историческом взгляде, в список могли и не попасть лучшие источники для получения дополнительных сведений.

Помимо линейного, существует много других важных методов криптоанализа. Кое-какие из них упомянуты в книге, но только тогда, когда мы понимали, как связать их с нашим изложением линейного криптоанализа. Поэтому некоторые важные криптоаналитические методы вообще не обсуждаются.

Как эту книгу может использовать начинающий

Эта книга основана на односеместровом курсе линейного криптоанализа, который мы впервые прочли в Лёвенском католическом университете осенью

2023 года. Это был первый курс криптоанализа, рассчитанный на студентов-магистрантов, знакомых с математикой и математическими методами. Этот учебник может лечь в основу похожих курсов, или же его можно прочитать от корки до корки в рамках самообразования. Впрочем, книгу можно читать и по частям, и некоторые рекомендации по этому поводу приведены ниже.

Главы 1–5 помогут освоить базовые принципы линейного криптоанализа, например их можно включить в более широкий курс криптоанализа. Главы 6–9 посвящены более специальным темам и могут быть полезны читателю, желающему углубить и расширить свои знания вплоть до современного состояния дел. В главах 10–11 как раз и обсуждается современное состояние дел; мы рекомендуем их тем, кто собирается изучать другие криптоаналитические методы, например дифференциальный и интегральный криптоанализ, а также исследователям, делающим свои первые самостоятельные шаги.

Для коротких курсов из одной-двух лекций мы не рекомендуем ограничиваться только главой 1. Эта глава поднимает больше вопросов, чем дает ответов. По той же причине начинающим следует быстро переходить к главе 2, а не пытаться понять каждое слово в главе 1 при первом чтении.

Читателям, больше интересующимся математическими аспектами линейного криптоанализа, нежели криптоанализом конкретных шифров, мы не рекомендуем долго задерживаться на главе 1, а главы 3, 5 и 9 они могут без опаски пропустить.

Как эту книгу может использовать специалист

Будучи сами исследователями и рецензируя чужие работы, мы иногда встречаем такие, где используются устаревшие методы и чрезмерно упрощенные аппроксимации. С помощью этой книги мы рассчитываем помочь в распространении знаний о современном состоянии дел. Специалисты, возможно, сочтут ее полезным справочником благодаря кое-какой актуальной информации, обзор которой приведен ниже.

Уже в главе 2 мы выдвигаем на первый план определение линейного криптоанализа с помощью корреляционных матриц. На наш взгляд, это самый эффективный способ получить основные результаты, не слишком увеличивая уровень абстракции. Рано или поздно, без корреляционных матриц все равно не обойтись, а введя их раньше, мы упростим переход к главам 10 и 11.

Обсуждение статистических аспектов линейного криптоанализа – тонкая материя. С одной стороны, явные формулы полезны для понимания главных факторов, влияющих на стоимость атаки. С другой стороны, в интересах точности желательно использовать как можно меньше упрощений. Мы старались соблюсти баланс, приводя замкнутые формулы там, где это можно сделать, не увязнув в технических деталях и всякий раз точно указывая, на какие упрощения пришлось пойти. Следует иметь в виду, что большинство существенных аппроксимаций в главах 4 и 7 связаны со статистическим моделированием реальности (стратегия выборки, зависимость корреляций от ключей, рандомизация с неправильным ключом и т. д.), а не с математическими вопросами типа скорости сходимости в предельной теореме.

Наше изложение многомерного линейного криптоанализа в главе 6 оригинально в том смысле, что не зависит от выбора базиса пространства масок.

Этот подход к многомерным линейным аппроксимациям далее развивается в главе 11.

В главе 11 вводится геометрический подход к криптоанализу с относительно конкретной точки зрения с упором на линейный криптоанализ и некоторые тесно связанные с ним методы. Более общая трактовка потребовала бы математической подготовки за пределами линейной алгебры. Тем не менее мы попытались согласовать изложение результатов и примеров с общей теорией, представление о которой дают несколько упражнений. Например, с самого начала мы настаиваем на различии между $\mathbb{C}[G]$ и $\mathbb{C}^G$ – но мы не обсуждаем структуру коалгебры и алгебры этих пространств.

Глава 1

Введение

Приложений криптографии множество, их легко встретить в повседневной жизни. Но эта книга не о приложениях, а о тех базовых строительных блоках, на которых зиждется их безопасность. Эти строительные блоки называются *криптографическими примитивами*, и наша книга является введением в анализ их безопасности. Вместо того чтобы рассматривать разнообразные методы на начальном уровне, мы займемся углубленным изучением одного семейства методов – линейного криптоанализа.

В этой главе рассматривается история вопроса, которая привела к открытию линейного криптоанализа. Плюсом такого описания «от Адама» является конкретность, но вообще-то оно не очень эффективно. Однако поднимает важные вопросы, изучаемые в последующих главах.

1.1. Криптографические примитивы

Принимая во внимание дискретную природу современной криптографии, большинство примитивов оперируют битовыми строками фиксированной длины. В этой книге множество битовых векторов длины n обозначается $\mathbb{F}_2^n$, где $\mathbb{F}_2$ – поле целых чисел по модулю 2. Самыми известными примитивами являются *блочные шифры*. Блочный шифр с размером блока n – это семейство обратимых функций, отображающих $\mathbb{F}_2^n$ в $\mathbb{F}_2^n$. Функция, принадлежащая этому семейству, обозначается E_k, где индекс k – обычно битовый вектор – называется *ключом*.

1.1.1. Анализ

В большинстве приложений блочных шифров ключ хранится в секрете. Таким образом, безопасность блочного шифра определяется тем, насколько противнику трудно узнать (*восстановить*) его ключ. Однако определение безопасности блочного шифра можно обобщить, и зачастую так и поступают. Например, если имеется возможность опрашивать блочный шифр, то можно говорить о том, насколько трудно понять (*различить*), взаимодействуем ли мы с шифром или с алгоритмом-пустышкой, который возвращает случайные результаты[1].

Трудность атаки включает несколько аспектов, начиная со свойств реализующего ее алгоритма: времени его работы, требований к памяти, степени параллелизма, вероятности успеха и т. д. Следует также принимать во внимание

[1] Результаты должны быть согласованы с тем, что шифр является перестановкой.

количество и тип требуемой информации. В случае атаки с известным открытым текстом доступны пары вход–выход для примитива – входы выбираются из известного распределения. В случае атаки с выбранным открытым текстом входы задаются атакующим.

Существует простая стратегия атаки, которая работает для любого блочного шифра: *исчерпывающий поиск ключа*. Для нее требуется несколько известных пар (открытый текст, шифртекст): (x_1,y_1), ..., (x_q, y_q). Далее в цикле перебираются все возможные значения ключа k и для каждого проверяется, верно ли, что $y_i = \mathsf{E}_k(x_i)$ для $i = 1, \ldots, q$. Исчерпывающий поиск ключа требует мало памяти и легко распараллеливается. Часто его используют как эталон для оценки релевантности других атак: чтобы алгоритм можно было квалифицировать как атаку, он должен превосходить исчерпывающий поиск хотя бы в одном аспекте.

1.1.2. Проектирование

Шифры можно конструировать путем композиции сравнительно простых функций:

$$\mathsf{E}_k = \mathrm{R}_k^{(r)} \circ \cdots \circ \mathrm{R}_k^{(1)}.$$

«Сравнительно простые» обычно означает, что функции $\mathrm{R}_k^{(r)}$, ..., $\mathrm{R}_k^{(1)}$ допускают эффективное вычисление на целевой платформе (платформах) и имеют компактное и хорошо понятное математическое описание. Все современные блочные шифры идут по этому пути.

Итеративные шифры – это шифры, в которых функции $\mathrm{R}_k^{(i)}$ являются экземплярами семейства функций с одним ключом:

$$\mathsf{E}_k = \mathrm{R}_{k_r} \circ \cdots \circ \mathrm{R}_{k_1}.$$

Функции R_{k_i} называются *раундами* F_k. Последовательность $(k_1, \ldots, k_r)$ называется *расширенным ключом* блочного шифра. Она строится путем применения функции, называемой *разверткой ключа*, к ключу k.

Шифры с чередованием ключа – это итеративные шифры, в которых раундовая функция является композицией функции, независимой от ключа, и прибавления ключа (в поле $\mathbb{F}_2^n$):

$$\mathrm{R}_{k_i}(x) = \mathrm{R}(x) + k_i.$$

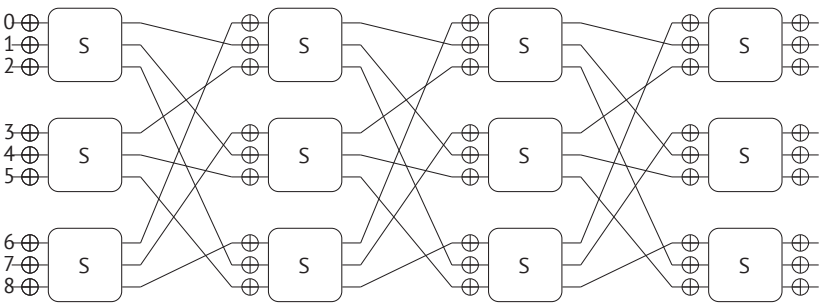


Рис. 1.1. Блочный шифр с размером блока 9 бит и четырьмя раундами

Термины «итеративный шифр» и «шифр с чередованием ключа» часто употребляются гибко: даже если первый или последний раунд немного отличается от прочих, шифр все равно называется итеративным или с чередованием ключа.

На рис. 1.1 изображен блочный шифр с размером блока $n = 9$. В этой главе он будет сквозным примером. Шифр представляет собой подстановочно-перестановочную сеть с ключом k длиной 45 бит, принадлежащим $\mathbb{F}_2^{45}$. Раундовая функция состоит из следующих трех операций.

S-блок. Эта операция применяет функцию $\mathsf{S}: \mathbb{F}_2^3 \to \mathbb{F}_2^3$ к трем группам битов состояния:

$$(x_8, \dots, x_0) \mapsto \mathsf{S}(x_8, x_7, x_6) \| \mathsf{S}(x_5, x_4, x_3) \| \mathsf{S}(x_2, x_1, x_0),$$

где символ «||» обозначает конкатенацию битовых векторов. S-блочная функция S впервые была использована в блочном шифре 3-Way и определена следующей таблицей подстановки. В приложении C приведен список всех упоминаемых в книге шифров, включая 3-Way.

x	000	001	010	011	100	101	110	111
$\mathsf{S}(x)$	111	010	100	101	001	110	011	000

Перестановка битов. Вторая операция переставляет биты состояния, отображая i-й выходной бит S-блока j на входной бит $j + 1 \pmod 3$ S-блока i. Конкретно $(x_8, \dots, x_0) \mapsto (x_5, x_2, x_8, x_4, x_1, x_7, x_3, x_0, x_6)$.

Сложение с ключом. Каждый раунд завершается прибавлением раундового ключа к состоянию. На i-м раунде (нумерация начинается с 1) операции сложения с ключом соответствует функция $(x_8, \dots, x_0) \mapsto (x_8 + k_{9i+8}, \dots, x_0 + k_{9i})$. На рис. 1.1 сложение с ключом представлено символом $\oplus$.

После прибавления битов ключа $(k_8, \dots, k_0)$ к открытому тексту шифр последовательно вычисляет эти операции четыре раза.

Во избежание недопонимания подчеркнем, что в этом примере использован учебный шифр, который на практике применять не следует. Из-за малого размера ключа (45 бит) становится возможен исчерпывающий поиск (поскольку число возможных ключей равно всего лишь 2^{45}) и даже более эффективные атаки, которые будут описаны ниже в этой главе. Также отметим, что в большинстве реальных шифров размер блока гораздо больше. Например, в шифре Advanced Encryption Standard (AES) он равен 128 бит.

1.2. Линейные аппроксимации

Линейный криптоанализ основан на *линейных аппроксимациях*. Это вероятностные линейные соотношения между входными и выходными битами функции $\mathsf{F}: \mathbb{F}_2^n \to \mathbb{F}_2^n$. Говоря «вероятностные», мы имеем в виду, что это соотношение имеет место не для всех входных значений функции. А под линейностью мы

понимаем линейность над полем $\mathbb{F}_2$. Если $y = \mathsf{F}(x)$, то линейной аппроксимации соответствует уравнение вида

$$\sum_{i=1}^{m} v_i\, y_i = \sum_{i=1}^{n} u_i x_i.$$

Его можно более компактно записать как $v^{\mathsf{T}}\mathsf{F}(x) = u^{\mathsf{T}}x$, где u и v – векторы с элементами $(u_1, \ldots, u_n)$ и $(v_1, \ldots, v_m)$ соответственно. Иногда мы рассматриваем u и v как битовые строки. Векторы u и v называются входной и выходной маской соответственно. Поскольку маски u и v определяют аппроксимацию, мы говорим, что линейная аппроксимация *является* парой масок (u, v) в пространстве $\mathbb{F}_2^n \times \mathbb{F}_2^m$.

1.2.1. Смещение

Пусть $\mathbf{x}$ – равномерно распределенная случайная величина, принимающая значения из $\mathbb{F}_2^n$. Рассмотрим вероятность линейной аппроксимации (u, v) функции F:

$$\Pr_{\mathbf{x}}\left[u^{\mathsf{T}}\mathbf{x} = v^{\mathsf{T}}\mathsf{F}(\mathbf{x})\right] = \frac{\left|\left\{x \in \mathbb{F}_2^n \mid u^{\mathsf{T}}x = v^{\mathsf{T}}\mathsf{F}(x)\right\}\right|}{2^n}.$$

Если вышеупомянутая вероятность равна $1/2$, то $u^{\mathsf{T}}x$ и $v^{\mathsf{T}}\mathsf{F}(x)$ никак не связаны: для половины входов x они принимают одинаковое значение, а для другой половины – дополнительные значения. Исходя из этого наблюдения, *смещение* $\epsilon_{u,v}$ линейной аппроксимации (u, v) функции F определяется как

$$\epsilon_{u,v} = \Pr_{\mathbf{x}}\left[u^{\mathsf{T}}\mathbf{x} = v^{\mathsf{T}}\mathsf{F}(\mathbf{x})\right] - \frac{1}{2}.$$

Если $\epsilon_{u,v} \neq 0$, то линейная аппроксимация (u, v) называется *эффективной*.

Пример 1.1. Пусть S – S-блок из демонстрационного шифра, определенного в разделе 1.1. Рассмотрим аппроксимацию $(u, v) = (001, 011)$ функции S. Для вычисления смещения построим следующую таблицу:

x	$u^{\mathsf{T}}x$	$\mathsf{S}(x)$	$v^{\mathsf{T}}\mathsf{S}(x)$
000	0	111	0
001	1	010	1
010	0	100	0
011	1	101	1
100	0	001	1
101	1	110	1
110	0	011	0
111	1	000	0

Отсюда следует, что смещение (001, 011) равно $6/8 - 1/2 = 1/4$. В качестве упражнения можете показать, что смещение аппроксимации (100, 100) равно $-1/4$. ▷

1.2.2. Таблицы линейной аппроксимации

Таблицей линейной аппроксимации (*linear approximation table* – LAT) функции $\mathsf{F} : \mathbb{F}_2^n \times \mathbb{F}_2^m$ называется таблица, содержащая смещения всех линейных аппроксимаций F, умноженная на масштабный коэффициент 2^n. То есть

$$\mathsf{LAT}_{u,v} = 2^n \epsilon_{u,v}.$$

С учетом нулевых масок всего существует 2^{n+m} аппроксимаций, и таблица LAT содержит 2^n строк и 2^m столбцов. Заметим, что элементы LAT индексированы битовыми векторами, т. е. элементами $\mathbb{F}_2^n$ и $\mathbb{F}_2^m$.

Теорема 1.1. *Пусть* F *– функция из* $\mathbb{F}_2^n$ *в* $\mathbb{F}_2^m$. *LAT* F *обладает следующими свойствами:*

1. $\mathsf{LAT}_{0,0} = 2^{n-1}$.
2. *Для всех ненулевых u, принадлежащих* $\mathbb{F}_2^n$, $\mathsf{LAT}_{u,0} = 0$.

Если F *обратима, то LAT* F *дополнительно обладает следующими свойствами:*

3. *Для всех ненулевых v, принадлежащих* $\mathbb{F}_2^m$, $\mathsf{LAT}_{v,0} = 0$.
4. *Все элементы LAT – четные числа.*

Доказательство. Первое свойство вытекает из того, что $\epsilon_{0,0} = \Pr_{\mathbf{x}}[0 = 0] - 1/2 = 1/2$. Что касается второго свойства, заметим, что

$$\epsilon_{u,0} = \frac{\left|\left\{x \in \mathbb{F}_2^n \mid u^{\mathsf{T}} x = 0\right\}\right|}{2^n} - \frac{1}{2}.$$

Для любого $u \neq 0$ существует 2^{n-1} значений x, принадлежащих $\mathbb{F}_2^n$, таких, что $u^{\mathsf{T}}x = 0$. Поэтому первый член в выражении выше равен $1/2$, и результат равен 0. Если F обратима, то $m = n$, и третье свойство доказывается аналогично. Действительно,

$$\epsilon_{0,v} = \frac{\left|\left\{x \in \mathbb{F}_2^n \mid v^{\mathsf{T}}\mathsf{F}(x) = 0\right\}\right|}{2^n} - \frac{1}{2} = \frac{\left|\left\{y \in \mathbb{F}_2^n \mid v^{\mathsf{T}} y = 0\right\}\right|}{2^n} - \frac{1}{2},$$

где второе равенство имеет место, потому что F обратима.

Если $u = 0$ или $v = 0$, то четвертое свойство следует из свойств (1)–(3). В противном случае обе функции $x \mapsto u^{\mathsf{T}}x$ и $x \mapsto v^{\mathsf{T}}\mathsf{F}(x)$ принимают значение 0 в точности для 2^{n-1} входов. Обозначим a количество значений x таких, что $u^{\mathsf{T}}x = 0$ и $v^{\mathsf{T}}\mathsf{F}(x) = 0$. Это приводит к следующему разбиению $\mathbb{F}_2^n$:

	$u^{\mathsf{T}}x = 0$	$u^{\mathsf{T}}x = 1$
$v^{\mathsf{T}}\mathsf{F}(x) = 0$	a	$2^{n-1} - a$
$v^{\mathsf{T}}\mathsf{F}(x) = 1$	$2^{n-1} - a$	a

В частности, существует $2^{n-1} - a$ значений x таких, что $u^{\mathsf{T}}x = 1$ и $v^{\mathsf{T}}\mathsf{F}(x) = 0$. Поскольку F обратима, существует также $2^{n-1} - a$ значений x таких, что $u^{\mathsf{T}}x = 0$ и $v^{\mathsf{T}}\mathsf{F}(x) = 1$. Отсюда следует, что существует $2^{n-1} - (2^{n-1} - a) = a$ значений x таких, что $u^{\mathsf{T}}x = 1$ и $v^{\mathsf{T}}\mathsf{F}(x) = 1$.

И наконец, количество x таких, что $u^{\mathsf{T}}x = v^{\mathsf{T}}\mathsf{F}(x)$, равно $2a$. □

Пример 1.2. Таблица линейной аппроксимации S равна

$$\mathsf{LAT} = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 2 & 0 & -2 & 0 & -2 \\ 0 & 0 & -2 & -2 & 0 & 0 & 2 & -2 \\ 0 & -2 & 2 & 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & -2 & 2 & -2 & -2 \\ 0 & 2 & 0 & 2 & -2 & 0 & 2 & 0 \\ 0 & 0 & -2 & 2 & 2 & 2 & 0 & 0 \\ 0 & -2 & -2 & 0 & -2 & 0 & 0 & 2 \end{bmatrix}.$$

В качестве упражнения проверьте свойства, перечисленные в теореме 1.1. ▷

Для некоторых функций LAT легко найти аналитически. Ниже приведено два примера – оба используются в демонстрационном шифре из раздела 1.1.

Сложение с константой. Пусть F – функция, которая прибавляет константу c к своему аргументу, т. е. $\mathsf{F}(x) = x + c$. Для всех линейных аппроксимаций (u, v) функции F имеем

$$\Pr_{\mathbf{x}}\left[u^{\mathsf{T}}\mathbf{x} = v^{\mathsf{T}}\mathsf{F}(\mathbf{x})\right] = \Pr_{\mathbf{x}}\left[u^{\mathsf{T}}\mathbf{x} = v^{\mathsf{T}}\mathbf{x} + v^{\mathsf{T}}c\right] = \Pr_{\mathbf{x}}\left[(u+v)^{\mathsf{T}}\mathbf{x} = v^{\mathsf{T}}c\right].$$

Если $u \neq v$, то вероятность равна 1/2 и, следовательно, смещение равно нулю. Если $u = v$, то вероятность равна единице, если $v^{\mathsf{T}}c = 0$, и нулю, если $v^{\mathsf{T}}c = 1$. Если принять соглашение, что $(-1)^b = 1$ для $b = 0$ в $\mathbb{F}_2$ и что $(-1)^b = -1$ для $b = 1$ в $\mathbb{F}_2$, то смещение равно

$$\epsilon_{u,v} = \begin{cases} (-1)^{v^{\mathsf{T}}c}\,\frac{1}{2}, & \text{если } u = v, \\ 0 & \text{в противном случае.} \end{cases}$$

Хотя ключ блочного шифра секретный, он все-таки является константой! Следовательно, линейная аппроксимация сложения с ключом описывается приведенной выше формулой.

Перестановка битов. Легко видеть, что если F – перестановка битов, то вероятность линейной аппроксимации (u, v) функции F равна единице, если $v = \mathsf{F}(u)$, и ½ в противном случае. То есть

$$\epsilon_{u,v} = \begin{cases} \frac{1}{2}, & \text{если } v = \mathsf{F}(u), \\ 0 & \text{в противном случае.} \end{cases}$$

1.3. Линейные следы и лемма о набегании знаков

В этом разделе мы займемся задачей о нахождении смещения линейной аппроксимации композиции функций $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_1$ в случае, когда известны только смещения линейных аппроксимаций функций $\mathsf{F}_1, \ldots, \mathsf{F}_r$. Этот вопрос относится прежде всего к анализу итеративных шифров.

Пусть $\mathbf{z}_1$ – равномерно распределенная случайная величина и $\mathbf{z}_{i+1} = \mathsf{F}_i(\mathbf{z}_i)$ для $i = 1, \ldots, r$. Чтобы найти смещение $\epsilon_{u_1,u_{r+1}}$ линейной аппроксимации (u_1, u_{r+1}) функции F, рассмотрим последовательные линейные аппроксимации функций $\mathsf{F}_1, \ldots, \mathsf{F}_r$ такие, что выходная маска каждой аппроксимации равна входной маске следующей. Последовательность масок $(u_1, \ldots, u_{r+1})$ называется *линейным следом*. Чтобы найти $\epsilon_{u_1,u_{r+1}}$, определим случайные величины $\mathbf{x}_1, \ldots, \mathbf{x}_r$ следующим образом:

$$\begin{aligned}
\mathbf{x}_1 &= u_1^\intercal \mathbf{z}_1 + u_2^\intercal \mathbf{z}_2 \\
\mathbf{x}_2 &= u_2^\intercal \mathbf{z}_2 + u_3^\intercal \mathbf{z}_3 \\
&\vdots \\
\mathbf{x}_r &= u_r^\intercal \mathbf{z}_r + u_{r+1}^\intercal \mathbf{z}_{r+1} \\
\hline
\textstyle\sum_{i=1}^r \mathbf{x}_r &= u_1^\intercal \mathbf{z}_1 + u_{r+1}^\intercal \mathbf{z}_{r+1}\,.
\end{aligned}$$

Смещение $\mathbf{x}_i$, т. е. $\Pr[\mathbf{x}_i = 0] - \frac{1}{2}$, равно смещению линейной аппроксимации (u_i, u_{i+1}) функции F_i. В общем случае смещение $\sum_{i=1}^r \mathbf{x}_i$ невозможно определить, зная смещения $\mathbf{x}_1, \ldots, \mathbf{x}_r$. Однако если $\mathbf{x}_1, \ldots, \mathbf{x}_r$ независимы, то его можно вычислить, воспользовавшись леммой о набегании знаков.

Лемма 1.2 *(о набегании знаков). Пусть $\mathbf{x}_1, \ldots, \mathbf{x}_r$ – случайные величины на $\mathbb{F}_2$ со смещениями $\epsilon_1, \ldots, \epsilon_r$. Если $\mathbf{x}_1, \ldots, \mathbf{x}_r$ независимы, то смещение ϵ суммы $\mathbf{x}_1 + \ldots + \mathbf{x}_r$ равно*

$$\epsilon = 2^{r-1} \prod_{i=1}^{r} \epsilon_i\,.$$

Доказательство. Рассмотрим случай $r = 2$. Смещение $\mathbf{x}_1 + \mathbf{x}_2$ удовлетворяет соотношению

$$\frac{1}{2} + \epsilon = \left(\frac{1}{2} + \epsilon_1\right)\left(\frac{1}{2} + \epsilon_2\right) + \left(1 - \frac{1}{2} - \epsilon_1\right)\left(1 - \frac{1}{2} - \epsilon_2\right).$$

Раскрывая скобки в правой части, получаем $\epsilon = 2\epsilon_1\epsilon_2$. Результат в общем случае получается рекурсивным применением формулы для $r = 2$. □

В случае линейного следа случайные величины $\mathbf{x}_1, \ldots, \mathbf{x}_r$, очевидно, не являются независимыми. Тем не менее лемма о набегании знаков используется как эвристика для оценки смещения линейной аппроксимации композиции функций:

$$\epsilon_{u_1,u_{r+1}} \approx 2^{r-1} \prod_{i=1}^{r} \epsilon_{u_i,u_{i+1}}.$$

Обсуждение точности этой эвристики прямо сейчас завело бы нас слишком далеко. Поэтому мы отложим его до главы 2, где формализм корреляционных матриц позволит решить этот вопрос просто.

Пример 1.3. (Линейный след для демонстрационного шифра.) Чтобы найти нетривиальную эффективную аппроксимацию для трех раундов демонстрационного шифра из раздела 1.1, можно скомбинировать три эффективные одно-раундовые аппроксимации.

Обозначим **a** (случайный равномерно распределенный) вход шифра, а **b**, **c** и **d** – входы первого, второго и третьего раундов. Наконец, пусть **e** – выход третьего раунда. Воспользовавшись LAT S (из примера 1.2) и нашими предыдущими наблюдениями для случаев сложения с константой и перестановки битов, можно проверить, что

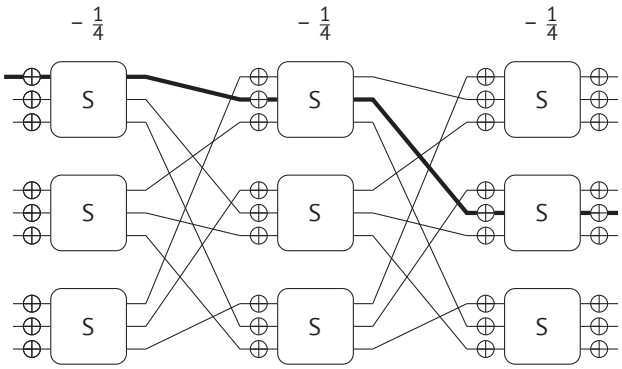


Рис. 1.2. Линейный след из примера 1.3

$\mathbf{a}_0 + \mathbf{b}_0 = 0$ со смещением $\epsilon_0 = \frac{1}{2}(-1)^{k_0}$,
$\mathbf{b}_0 + \mathbf{c}_1 = 0$ со смещением $\epsilon_1 = -\frac{1}{4}(-1)^{k_{10}}$,
$\mathbf{c}_1 + \mathbf{d}_4 = 0$ со смещением $\epsilon_2 = -\frac{1}{4}(-1)^{k_{22}}$,
$\mathbf{d}_1 + \mathbf{e}_4 = 0$ со смещением $\epsilon_3 = -\frac{1}{4}(-1)^{k_{31}}$.

Маски, соответствующие этому следу, показаны на рис. 1.2, где жирными линиями обозначены ненулевые биты масок. Например, маска на входе третьего раунда равна 000010000.

Эвристически, в силу леммы о набегании знаков и тождества $(-1)^x (-1)^y = (-1)^{x+y}$, линейная аппроксимация (000000001, 000010000), или эквивалентно $\mathbf{a}_0 + \mathbf{e}_4 = 0$, имеет смещение

$$\epsilon \approx (-1)^{k_0+k_{10}+k_{22}+k_{31}+1} \frac{1}{16}.$$

В качестве упражнения попробуйте найти еще хотя бы один след с такими же входными и выходными масками и покажите, что абсолютная величина

его смещения меньше $1/16$. Так как разные следы дают разные результаты, лемму 1.2 следует использовать только для следа, который имеет наибольшее смещение. Однако, как показано в главе 2, даже в этом случае нет гарантии, что результаты точны. ▷

1.4. Восстановление ключа

Эффективную линейную аппроксимацию блочного шифра можно использовать для организации атаки с восстановлением ключа. Есть два основных способа сделать это: «алгоритм Мацуи 1» и «алгоритм Мацуи 2», названные в честь автора.

Оба метода полагаются на оценку смещения линейной аппроксимации по случайной выборке данных. Именно поэтому линейный криптоанализ часто называют статистической атакой.

1.4.1. Алгоритм Мацуи 1

Алгоритм Мацуи 1 восстанавливает 1 бит информации о расширенном ключе шифра с чередованием ключа[1].

Рассмотрим блочный шифр с чередованием ключа $\mathsf{E}_k = \mathsf{R}_{k_r} \circ \cdots \circ \mathsf{R}_{k_1}$, где $\mathsf{R}_{k_i}(x) = \mathsf{R}(x) + k_i$, и обозначим $\epsilon_{u_i,u^{i+1}}$ смещение линейной аппроксимации (u_i, u_{i+1}) функции R. Тогда смещение линейной аппроксимации (u_i, u_{i+1}) функции R_{k_i} равно

$$(-1)^{u_{i+1}^{\mathsf{T}} k_i} \epsilon_{u_i, u_{i+1}}.$$

В разделе 1.3 смещение линейной аппроксимации оценивалось с помощью линейного следа. Пусть $(u_1, \ldots, u_{r+1})$ – линейный след композиции $\mathsf{E}_k = \mathsf{R}_{k_r} \circ \cdots \circ \mathsf{R}_{k1}$. Применив лемму о набегании знаков к этому следу, мы получим следующую оценку смещения аппроксимации (u_1, u_{r+1}):

$$\epsilon_{u_1, u_{r+1}} \approx 2^{r-1} \prod_{i=1}^{r} (-1)^{u_{i+1}^{\mathsf{T}} k_i} \epsilon_{u_i, u_{i+1}} = 2^{r-1} (-1)^{z} \prod_{i=1}^{r} \epsilon_{u_i, u_{i+1}}.$$

В правой части z равно $\sum_{i=1}^{r} u^{\mathsf{T}}_{i+1} k_i$. Это будет тот бит информации о секретном ключе, который восстанавливает алгоритм Мацуи 1. Заметим, что это не бит ключа в строгом смысле слова, а линейное выражение от нескольких битов расширенного ключа.

При условии, что аппроксимативная природа уравнения не изменяет знака, z можно вычислить по знаку $\prod_{i=1}^{r} \epsilon_{u_i,u_{i+1}}$ и $\epsilon_{u_1,u_{r+1}}$. Первый можно определить с помощью теоретического анализа следа. Наиболее вероятное значение второго получается из *эмпирического смещения* линейной аппроксимации (u_1, u_{r+1}). Эмпирическое смещение оценивает по случайной выборке пар (открытый текст, шифртекст).

---

[1] Алгоритм 1 имеет более широкое применение, но мы здесь ограничимся только случаем с чередованием ключа.

Имея случайную выборку q пар (открытый текст, шифртекст) $(\mathbf{x}_i, \mathbf{y}_i)$, мы вычисляем эмпирическое смещение линейной аппроксимации (u_1, u_{r+1}) по формуле

$$\widehat{\epsilon} = \frac{1}{q}\left|\left\{1 \le i \le q \mid u_1^{\mathsf{T}}\mathbf{x}_i = u_{r+1}^{\mathsf{T}}\mathbf{y}_i\right\}\right| - \frac{1}{2}.$$

Среднее $\hat{\epsilon}$ равно $\epsilon_{u_1,u_{r+1}}$. Для независимых выборок дисперсия количества случаев, когда аппроксимация имеет место, близка к $q/4$. Действительно, для одной выборки дисперсия равна $(\frac{1}{2} + \epsilon)(1 - \frac{1}{2} - \epsilon) \approx \frac{1}{4}$. Следовательно, стандартное отклонение ϵ приближенно равно $1/\sqrt{4q}$. Отсюда следует, что для определения знака $\epsilon_{u_1,u_{r+1}}$ с высокой степенью достоверности требуется число образцов q такое, что

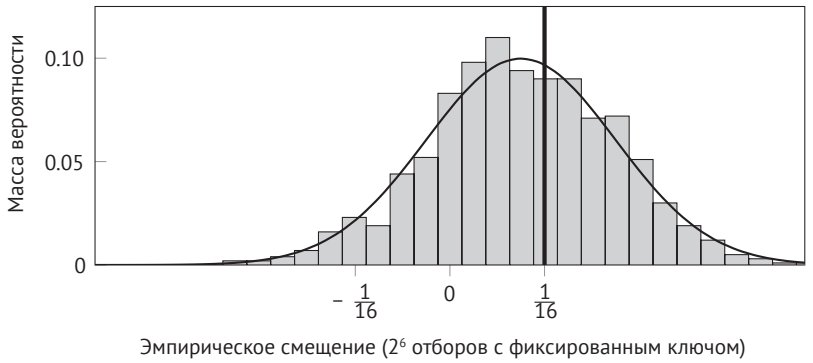


Рис. 1.3. Гистограмма эмпирического смещения для 1000 экспериментов

$$\frac{1}{\sqrt{4q}} \ll |\epsilon_{u_1,u_{r+1}}|.$$

Таким образом, для алгоритма Мацуи 1 требуется $q \gg 1/(2\epsilon_{u_1,u_{r+1}})^2$ образцов.

Пример 1.4 (алгоритм Мацуи 1 для демонстрационного шифра). Рассмотрим линейную аппроксимацию из примера 1.3. С помощью линейного следа мы оценили его смещение как $-\frac{1}{16} \cdot (-1)^z$, где $z = k_0 + k_{10} + k_{22} + k_{31}$. Поэтому 64-х образцов должно быть достаточно для определения z. Чтобы проверить, насколько хорошо работает эта атака, мы выполнили ее 1000 раз (с ключом `000000001 010000000 000000000 000000000`) и вычислили эмпирическое смещение. Гистограмма результатов показана на рис. 1.3.

Среднее эмпирическое смещение для 1000 экспериментов оказалось чуть меньше $\frac{1}{16}$. Это не совпадение – из результатов главы 2 следует, что в действительности для использованного в эксперименте ключа смещение равно $\frac{3}{64}$. Простое взятие знака эмпирической корреляции, скорее всего, даст $z = 0$ (правильное значение). ▷

Линейный криптоанализ обычно называется атакой с известным открытым текстом, но заметим, что для применения алгоритма Мацуи 1 полные открытые и шифртексты не нужны, достаточно значений $u_1^{\mathsf{T}}\mathbf{x}_i$ и $u_{r+1}^{\mathsf{T}}\mathbf{y}_i$. В упражнении 1.6 вам

будет предложено доказать, что это наблюдение позволяет обобщить алгоритм Мацуи 1 на случай, когда известна только оценка $\Pr_{\mathbf{x}_i}[u_1^{\mathsf{T}}\mathbf{x}_i = 0]$ (помимо $u_{r+1}^{\mathsf{T}}\mathsf{F}(\mathbf{x})$). Это наблюдение используется также в алгоритме Мацуи 2, описанном в разделе 1.4.2.

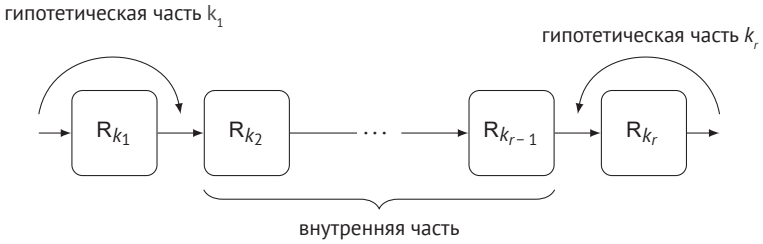


Рис. 1.4. Разбиение итеративного шифра на внутреннюю и внешнюю части

1.4.2. Алгоритм Мацуи 2

Алгоритм Мацуи 2 разбивает блочный шифр на две части, как показано на рис. 1.4.

Внешняя часть, где большие части раундовых ключей восстанавливаются путем угадывания.

Внутренняя часть, где линейная аппроксимация применяется для фильтрации гипотез о раундовых ключах, сделанных во внешней части, и где для восстановления линейного выражения от некоторых битов расширенного ключа можно использовать алгоритм Мацуи 1.

Для простоты мы опишем алгоритм для случая, когда внешняя часть состоит только из последнего раунда. В общем случае можно включить несколько раундов в начало или конец шифра при условии, что требуется угадать не слишком много битов ключа.

Если обозначить G_k внутреннюю часть, то $\mathsf{E}_k = \mathsf{R}_{k_r} \circ \mathsf{G}_k$, откуда $\mathsf{G}_k = \mathsf{R}_{k_r}^{-1} \circ \mathsf{E}_k$. Для линейной аппроксимации (u_1, u_r) функции G_k эмпирическая корреляция равна

$$\widehat{\epsilon} = \frac{1}{q}\left|\left\{1 \le i \le q \mid u_1^{\mathsf{T}}\mathbf{x}_i = u_r^{\mathsf{T}}\mathsf{R}_{k_r}^{-1}(\mathbf{y}_i)\right\}\right| - \frac{1}{2}.$$

Поскольку k_r заранее неизвестно, мы не можем определить $u_r^{\mathsf{T}}\mathsf{R}_{k_r}^{-1}(y)$ по y. Однако типичная функция R_{k_r} обладает тем свойством, что для некоторых масок u_r для вычисления $u_r^{\mathsf{T}}\mathsf{R}_{k_r}^{-1}(y)$ по y необходимо лишь небольшое число битов раундового ключа k_r.

Далее алгоритм Мацуи 2 оценивает эмпирическое смещение для каждой гипотезы о ключе, используя одну и ту же случайную выборку пар (открытый текст, шифртекст). Предполагается, что для неверной гипотезы эмпирическое смещение близко к нулю – или по крайней мере гораздо ближе к нулю, чем для истинного значения ключа. Это предположение часто применяется на практике, хотя есть случаи, когда ряд «эквивалентных» гипотез о ключе дают сравнимые эмпирические смещения.

Алгоритм Мацуи 2 выводит те гипотетические ключи, для которых эмпирическое смещение дальше всего отстоит от нуля. Остальные гипотезы называ-

ются ключами-кандидатами. Таким образом, алгоритм Мацуи 2 дает больше информации о секретном ключе, чем алгоритм Мацуи 1.

Определение частоты успехов алгоритма Мацуи 2 прямо сейчас завело бы нас слишком далеко в сторону. Пока просто констатируем, что, как и в случае алгоритма Мацуи 1, объем данных, необходимый для достижения высокой частоты успехов, пропорционален $1/\epsilon^2_{u_1,u_r}$. В общем случае $\epsilon_{u_1,u_r} \geqslant \epsilon^2_{u_1,u_{r+1}}$. Поэтому алгоритму Мацуи 2 может понадобиться меньше данных, чем алгоритму Мацуи 1. Однако частота успехов алгоритма Мацуи 2 зависит также от числа K значений ключа, которые считаются возможными априори, и от числа значений-кандидатов, возвращаемых в качестве выходов. Более детальный анализ приведен в главе 4.

При наивной реализации на qK вычислений $y \mapsto u_r^{\mathsf{T}} \mathsf{R}_{k_r}^{-1}(y)$ тратится преобладающая часть времени работы алгоритма Мацуи 2. Далее и, в частности, в главе 5 обсуждаются более быстрые способы вычисления эмпирических смещений.

1.5. Оставшиеся проблемы

В конце первой главы уместно будет упомянуть некоторые проблемы, которые мы до сих пор игнорировали. Сейчас вы уже знакомы с основной идеей линейного криптоанализа. Однако если бы вам пришлось применить полученные знания к атаке на реальные блочные шифры – или даже на демонстрационный шифр из раздела 1.1, – то вы, скорее всего, столкнулись бы с трудностями.

В разделе 1.3 мы использовали лемму о набегании знаков для оценки смещения линейной аппроксимации композиции функций. Понимание точности этой оценки составляет важную часть главы 2.

Другой вопрос – как найти линейные аппроксимации и линейные следы шифра с наибольшим (по абсолютной величине) смещением. Эта проблема обсуждается в главе 3.

Наконец, в нашем обсуждении атак с восстановлением ключа в разделе 1.4 игнорируются такие важные аспекты, как вероятность успеха описанных методов. Важно хорошо понимать, сколько данных потребуется для восстановления ключа. Этот вопрос подробно обсуждается в главе 4.

1.6. Историческая справка

Линейные аппроксимации и их смещение тесно связаны с другими концепциями, которые уже использовались ранее для анализа булевых функций, например с преобразованием Уолша–Адамара и минимальным расстоянием Хэмминга до аффинной функции. В упражнении 1 исследуется эта последняя идея. Несмотря на то что эти концепции изучались в контексте криптоанализа, ключевые составные части линейного криптоанализа отсутствовали.

Впервые линейные аппроксимации применили в криптоанализе Анна Тарди-Корфдир и Анри Жильбер в 1991 году. Они использовали линейные аппроксимации частей блочного шифра FEAL для организации атаки с восстановлением ключа. Термины «линейный криптоанализ» и «лемма о набегании значений» ввел Мацуи в 1993 году. Он использовал линейный криптоанализ для атаки на блочный шифр *Data Encryption Standard* (DES).

1.7. Литература

Matsui, Mitsuru (May 1994a). «Linear Cryptanalysis Method for DES Cipher». In: *EUROCRYPT'93*. Ed. by Tor Helleseth. Vol. 765. LNCS. Springer, Berlin, Heidelberg, pp. 386–397. doi: 10.1007/3-540-48285-7_33.

Matsui, Mitsuru (Aug. 1994b). «The First Experimental Cryptanalysis of the Data Encryption Standard». In: *CRYPTO'94*. Ed. by Yvo Desmedt. Vol. 839. LNCS. Springer, Berlin, Heidelberg, pp. 1–11. doi: 10.1007/3-540-48658-5_1.

Tardy-Corfdir, Anne and Henri Gilbert (Aug. 1992). «A Known Plaintext Attack of FEAL-4 and FEAL-6». In: *CRYPTO'91*. Ed. by Joan Feigenbaum. Vol. 576. LNCS. Springer, Berlin, Heidelberg, pp. 172–181. doi: 10.1007/3-540-46766-1_12.

1.8. Упражнения

Упражнение 1.1

Пусть $\mathsf{F}(x) = k_2 + \mathsf{S}(k_1 + x)$, где k_1, k_2 – ключи, а S – S-блок, показанный в табл. 1.1.

1. Найдите нетривиальную линейную аппроксимацию F.
2. Примените алгоритм Мацуи 1 для восстановления одного бита ключа.

Таблица 1.1. 4-битовый S-блок S, значения записаны в шестнадцатеричном виде (например, e = 1110)

0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
8	2	4	0	f	5	7	c	a	6	b	3	e	d	9	1

Упражнение 1.2

Функция $f : \mathbb{F}_2^n \to \mathbb{F}_2$ называется линейной, если $f(00 \ldots 0) = 0$ и $f(x + y) = f(x) + f(y)$ для всех $x, y \in \mathbb{F}_2^n$.

Для любого $u \in \mathbb{F}_2^n$ обозначим $\ell_u : \mathbb{F}_2^n \to \mathbb{F}_2$ линейную функцию, определенную как $\ell_u(x) = u^\mathsf{T} x$.

1. Покажите, что для любой линейной функции $f : \mathbb{F}_2^n \to \mathbb{F}_2$ существует маска $u \in \mathbb{F}_2^n$ такая, что $f = \ell_u$.
2. Постройте таблицы истинности всех 3-битовых линейных функций $\mathbb{F}_2^3 \to \mathbb{F}_2$.

Упражнение 1.3

Еще до появления линейного криптоанализа было известно, что S-блок $\mathsf{S}_5 : \mathbb{F}_2^6 \to \mathbb{F}_2^4$ шифра DES (см. табл. 1.2) обладает специальным «линейным свойством».

1. Вычислите LAT блока S_5.
2. Что такое специальное линейное свойство?

Таблица 1.2. Шестнадцатеричное табличное представление S-блока S_5 шифра DES

2	e	c	b	4	2	1	c	7	4	a	7	b	d	6	1	...
8	5	5	0	3	f	f	a	d	3	0	9	e	8	9	6	...
4	b	2	8	1	c	b	7	a	1	d	e	7	2	8	d	...
f	6	9	f	c	0	5	9	6	a	3	4	0	5	e	3	

Упражнение 1.4

Пусть $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^{2n}$– операция n-битового разветвления, определенная как $\mathsf{F}(x) = x\|x$. Вычислите LAT функции F.

Упражнение 1.5

Определим расстояние Хэмминга между функциями $f : \mathbb{F}_2^n \to \mathbb{F}_2$ и $g : \mathbb{F}_2^n \to \mathbb{F}_2$ как $d_H(f, g) = \mathrm{wt}(f + g)$, где wt – вес Хэмминга (число единиц) таблицы истинности $f + g$.

1. Нелинейность функции $f : \mathbb{F}_2^n \to \mathbb{F}_2$ определяется как
$$\mathcal{N}(f) = \min_{\substack{u \in \mathbb{F}_2^n \\ a \in \mathbb{F}_2}} d_H(f, \ell_u + a).$$
Докажите, что $\mathcal{N}(f) = 2^{n-1} - \max_{u \in \mathbb{F}_2^n} \left| |\{x \in \mathbb{F}_2^n \mid f(x) = \ell_u(x)\}| - 2^{n-1} \right|$.
О том, что такое ℓ_u, см. упражнение 1.2.
2. Нелинейность функции $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ определяется как
$$\mathcal{N}(\mathsf{F}) = \min_{v \in \mathbb{F}_2^m \setminus \{0\}} \mathcal{N}(\ell_v \circ \mathsf{F}).$$

Докажите, что
$$\mathcal{N}(\mathsf{F}) = 2^{n-1} - \max_{\substack{u \in \mathbb{F}_2^n \\ v \in \mathbb{F}_2^m \setminus \{0\}}} |\mathsf{LAT}^{\mathsf{F}}_{u,v}|.$$

Упражнение 1.6

1. Обобщите алгоритм Мацуи 1 на случай, когда, помимо $u_{r+1}^{\mathsf{T}}\mathsf{F}(x)$, известна только оценка вероятности $\Pr_{\mathbf{x}}[u_1^{\mathsf{T}} = 0]$.
2. Как бы вы использовали это обобщение, если бы открытым текстом являлся английский текст в кодировке UTF-8?

Упражнение 1.7

Покажите, что, тщательно выбирая входы, смещение линейной аппроксимации обратимой функции можно найти, вычислив функцию только в половине входов.

Глава 2

Корреляционные матрицы

В главе 1 мы оценивали корреляции линейных аппроксимаций с помощью поиска подходящего линейного следа и применения леммы о набегании знаков, но этот подход опирался на ничем не обоснованное предположение о независимости. В этой главе под лемму о набегании знаков и линейный криптоанализ вообще подводится более солидный теоретический фундамент. Достигается это с помощью теории *корреляционных матриц*. Дамен предложил эти матрицы в 1994 году, чтобы упростить описание линейного криптоанализа.

2.1. Корреляция случайной величины на $\mathbb{F}_2$

Напомним (см. главу 1), что смещение линейной аппроксимации равно заключенной в ней вероятности минус одна вторая. На протяжении всей главы 1 и в разделе 1.3 в частности термин «смещение» употреблялся также в более общем смысле по отношению к случайному биту $\mathbf{x}$, т. е. случайной величине на $\mathbb{F}_2$. Точнее, смещение $\mathbf{x}$ равно

$$\epsilon_{\mathbf{x}} = \Pr_{\mathbf{x}}[\mathbf{x} = 0] - \frac{1}{2}.$$

В этой главе будет показано, что *корреляция* $\mathbf{x}$ – величина, с которой работать более естественно. Корреляция $\mathbf{x}$ просто равна удвоенному смещению:

$$c_{\mathbf{x}} = 2\epsilon_{\mathbf{x}} = 2\Pr_{\mathbf{x}}[\mathbf{x} = 0] - 1.$$

Если $\mathbb{E}\mathbf{X}$ обозначает среднее случайной величины $\mathbf{X}$, то корреляцию случайного бита $\mathbf{x}$ можно также записать в виде

$$c_{\mathbf{x}} = \Pr_{\mathbf{x}}[\mathbf{x} = 0] - \Pr_{\mathbf{x}}[\mathbf{x} = 1] = \mathbb{E}\,(-1)^{\mathbf{x}},$$

поскольку $(-1)^0 = 1$ и $(-1)^1 = -1$.

Еще одним основанием предпочесть корреляции смещениям служит то, что они упрощают формулировку и доказательство леммы о набегании знаков.

Лемма 2.1 *(о набегании знаков с корреляциями). Пусть* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r$ *– случайные величины на* $\mathbb{F}_2$. *Если* $\mathbf{x}_1, \ldots, \mathbf{x}_r$ *независимы, то корреляция суммы* $\mathbf{x}_1 + \ldots + \mathbf{x}_r$ *удовлетворяет соотношению*

$$c_{\mathbf{x}_1+\cdots+\mathbf{x}_r} = \prod_{i=1}^{r} c_{\mathbf{x}_i}.$$

Доказательство. Если $\mathbf{X}$ и $\mathbf{Y}$ – независимые случайные величины, то $\mathbb{E}\mathbf{XY} = (\mathbb{E}\mathbf{X})(\mathbb{E}\mathbf{Y})$. Отсюда

$$c_{\mathbf{x}_1+\cdots+\mathbf{x}_r} = \mathbb{E}\,(-1)^{\mathbf{x}_1+\cdots+\mathbf{x}_r} = \mathbb{E}\prod_{i=1}^{r}(-1)^{\mathbf{x}_i} = \prod_{i=1}^{r}\mathbb{E}\,(-1)^{\mathbf{x}_i}.$$

Второе равенство следует из того, что $(-1)^{x+y} = (-1)^x(-1)^y$ для любых $x, y \in \mathbb{F}_2$, что можно проверить отдельно для каждого случая. □

По сравнению с леммой 1.2 из главы 1, в лемме 2.1 отсутствует дополнительная степень двойки. Это удобно, потому что нам не нужно отслеживать количество компонуемых функций при применении леммы о набегании знаков в контексте линейного криптоанализа.

Замечание 2.2. Смысл использования корреляций не сводится к удобству обозначений: корреляция $c_{\mathbf{x}}$ – это нетривиальный коэффициент преобразования Фурье функции массы вероятности $\mathbf{x}$. На самом деле лемма о набегании знаков – это частный случай теоремы о свертке для преобразований Фурье. В последующих главах мы объясним эту связь лучше, чем возможно в настоящий момент. ▷

2.2. Корреляция между булевыми функциями

Помимо корреляции случайной величины на $\mathbb{F}_2$, существует родственное понятие корреляции между двумя булевыми функциями, которое иногда бывает полезно. Точнее, корреляция *между* функциями $f : \mathbb{F}_2^n \to \mathbb{F}_2$ и $g : \mathbb{F}_2^n \to \mathbb{F}_2$ определяется следующим образом:

$$\mathsf{C}(f,g) = 2\Pr_{\mathbf{x}}[f(\mathbf{x}) = g(\mathbf{x})] - 1,$$

где $\mathbf{x}$ – случайная равномерно распределенная величина на $\mathbb{F}_2^n$. Это не что иное, как корреляция случайной величины $f(\mathbf{x}) + g(\mathbf{x})$. По-другому $\mathsf{C}(f, g)$ называется *коэффициентом корреляции* между f и g.

Корреляция между двумя булевыми функциями – это мера их сходства. Если f и g равны, то $\mathsf{C}(f, g) = 1$. Если f и g всегда различны, то $\mathsf{C}(f, g) = -1$. Наконец, $\mathsf{C}(f, g) = 0$, если f и g равны на половине входов.

Пользуясь свойствами корреляций случайных величин, можно показать, что корреляция между булевыми функциями f и g равна (см. упражнение 2.1)

$$\mathsf{C}(f,g) = \frac{1}{2^n}\sum_{x\in\mathbb{F}_2^n}(-1)^{f(x)+g(x)} = \left\langle (-1)^f, (-1)^g \right\rangle,$$

где $\langle\cdot,\cdot\rangle$ – скалярное произведение функций $\mathbb{F}_2^n \to \mathbb{R}$. Следовательно, корреляции можно интерпретировать как скалярные произведения, что и является обоснованием термина «корреляция».

Любая линейная булева функция имеет вид $\ell_u(x) = u^\mathsf{T}x$ для некоторого $u \in \mathbb{F}_2^n$ (см. упражнение 1.2). Корреляция линейной аппроксимации (u, v) с масками $u \in \mathbb{F}_2^n$ и $v \in \mathbb{F}_2^m$ функции $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ равна $\mathsf{C}(\ell_u, \ell_v \circ \mathsf{F})$. Это действительно удвоенное смещение $\epsilon_{u,v}$ аппроксимации (u, v).

2.3. Корреляционные матрицы

В разделе 1.2.2 была определена таблица линейной аппроксимации (LAT) функции. Ниже определяется аналогичная таблица, содержащая корреляции всех линейных аппроксимаций функции.

Определение 2.3 (корреляционная матрица). Корреляционной матрицей функции $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ называется вещественная матрица C^F размера $2^m \times 2^n$ с элементами

$$C^\mathsf{F}_{v,u} = \mathsf{C}(\ell_u, \ell_v \circ \mathsf{F}) = \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{v^\mathsf{T}\mathsf{F}(x) + u^\mathsf{T}x}.$$

Корреляционная матрица F тесно связана с LAT F:

$$\mathsf{LAT}_{u,v} = 2^{n-1} C^\mathsf{F}_{v,u}.$$

Обратите внимание на порядок индексов: выходная маска – это индекс строки в корреляционных матрицах, но индекс столбца в LAT. В отличие от LAT, корреляционная матрица – больше, чем таблица, содержащая корреляции всех линейных аппроксимаций. Она представляет линейный оператор между вещественными векторными пространствами размерностей 2^n и 2^m. В описанных ниже свойствах корреляционных матриц этот факт уже предполагается известным, но с полным объяснением придется подождать до главы 11. У такого подхода есть недостаток – некоторые свойства могут показаться каким-то чудом, но так мы следуем историческому развитию предмета и оставляем эту главу более конкретной.

Теорема 2.4. *Пусть* $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ *и* $\mathsf{G} : \mathbb{F}_2^m \to \mathbb{F}_2^l$ *– функции с корреляционными матрицами* C^F *и* C^G *соответственно. Корреляционная матрица их композиции* $\mathsf{G} \circ \mathsf{F}$ *равна* $C^{\mathsf{G}\circ\mathsf{F}} = C^\mathsf{G} C^\mathsf{F}$.

Доказательство. По определению произведения двух матриц, элемент $(C^\mathsf{G} C^\mathsf{F})_{v,u}$ равен

$$\begin{aligned}\sum_{w \in \mathbb{F}_2^m} C^\mathsf{G}_{v,w} C^\mathsf{F}_{w,u} &= \sum_{w \in \mathbb{F}_2^m} \frac{1}{2^n} \frac{1}{2^m} \sum_{x \in \mathbb{F}_2^n} \sum_{y \in \mathbb{F}_2^m} (-1)^{v^\mathsf{T}\mathsf{G}(y) + w^\mathsf{T}y + w^\mathsf{T}\mathsf{F}(y) + u^\mathsf{T}x} \\ &= \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} \sum_{y \in \mathbb{F}_2^m} (-1)^{v^\mathsf{T}\mathsf{G}(y) + u^\mathsf{T}x} \frac{1}{2^m} \sum_{w \in \mathbb{F}_2^m} (-1)^{w^\mathsf{T}(y + \mathsf{F}(x))}.\end{aligned}$$

Правую часть можно переписать в виде

$$\sum_{w\in\mathbb{F}_2^m} C^{\mathsf{G}}_{v,w} C^{\mathsf{F}}_{w,u} = \frac{1}{2^n} \sum_{x\in\mathbb{F}_2^n} \sum_{y\in\mathbb{F}_2^m} (-1)^{v^{\mathsf{T}}\mathsf{G}(y)+u^{\mathsf{T}}x}\, \delta^y\big(\mathsf{F}(x)\big)$$
$$= \frac{1}{2^n} \sum_{x\in\mathbb{F}_2^n} (-1)^{v^{\mathsf{T}}\mathsf{G}(\mathsf{F}(x))+u^{\mathsf{T}}x}.$$

Здесь функция $\delta^y : \mathbb{F}_2^m \to \mathbb{R}$ определена как $\delta^y(y) = 1$ и $\delta^y(z) = 0$ для всех $z \neq y$. Выражение в последней строке в точности равно $C^{\mathsf{G}\circ\mathsf{F}}_{v,u}$. На первом шаге используется равенство

$$\sum_{w\in\mathbb{F}_2^m} (-1)^{w^{\mathsf{T}}a} = \begin{cases} 2^m, & \text{если } a = 0, \\ 0 & \text{в противном случае.} \end{cases}$$

Это следует из того, что для всех $t \in \mathbb{F}_2^m$ имеем

$$\sum_{w\in\mathbb{F}_2^m} (-1)^{w^{\mathsf{T}}a} = \sum_{w\in\mathbb{F}_2^m} (-1)^{(w+t)^{\mathsf{T}}a} = (-1)^{t^{\mathsf{T}}a} \sum_{w\in\mathbb{F}_2^m} (-1)^{w^{\mathsf{T}}a}.$$

Для любого ненулевого a существует по меньшей мере одно t такое, что $t^{\mathsf{T}}a = 1$. Следовательно, сумма противоположна самой себе, а значит, должна быть равна нулю. Если $a = 0$, то все члены суммы равны 1, так что она равна 2^m. □

Теорема 2.4 – самый важный результат, относящийся к корреляционным матрицам. Теоретически она дает способ вычислить корреляции линейных аппроксимаций $\mathsf{G} \circ \mathsf{F}$, зная только корреляции линейных аппроксимаций F и G. В главе 1 для достижения того же результата использовалась лемма о набегании знаков, но, в отличие от последней, теорема 2.4 не требует предположения о независимости.

Напомним, что матрица называется ортогональной, если обратная к ней совпадает с транспонированной. Следующий результат показывает, что корреляционные матрицы обратимых функций ортогональны.

Теорема 2.5. *Пусть* $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^n$ *– функция с корреляционной матрицей* C^{F}. *Если* F *является перестановкой, то* C^{F} *– ортогональная матрица.*

Доказательство. Нетрудно видеть, что корреляционная матрица тождественной функции является единичной (это также следует из теоремы 2.6). Следовательно, достаточно показать, что $C^{\mathsf{F}^{-1}} = (C^{\mathsf{F}})^{\mathsf{T}}$. Если F является перестановкой, то

$$C^{\mathsf{F}}_{v,u} = \frac{1}{2^n} \sum_{x\in\mathbb{F}_2^n} (-1)^{v^{\mathsf{T}}\mathsf{F}(x)+u^{\mathsf{T}}x} = \frac{1}{2^n} \sum_{y\in\mathbb{F}_2^n} (-1)^{v^{\mathsf{T}}y+u^{\mathsf{T}}\mathsf{F}^{-1}(y)} = C^{\mathsf{F}^{-1}}_{u,v}.$$

Существует другое доказательство, основанное на вычислении $(C^{\mathsf{F}})^{\mathsf{T}}C^{\mathsf{F}}$. □

Пример 2.1 (корреляционная матрица). Корреляционная матрица S-блока $\mathsf{S} : \mathbb{F}_2^3 \to \mathbb{F}_2^3$ демонстрационного шифра из раздела 1.1 равна

$$C^{\mathsf{S}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1/2 & 0 & -1/2 & 0 & 1/2 & 0 & -1/2 \\ 0 & 0 & -1/2 & 1/2 & 0 & 0 & -1/2 & -1/2 \\ 0 & 1/2 & -1/2 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & -1/2 & -1/2 & 1/2 & -1/2 \\ 0 & -1/2 & 0 & 1/2 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 & -1/2 & 1/2 & 0 & 0 \\ 0 & -1/2 & -1/2 & 0 & -1/2 & 0 & 0 & 1/2 \end{bmatrix}.$$

Элементы корреляционной матрицы индексированы битовыми векторами, поэтому чтобы представить ее в виде массива чисел, необходимо выбрать какой-нибудь произвольный порядок. Здесь выбран лексикографический порядок: $001 \leqslant 010 \leqslant 011 \leqslant \ldots \leqslant 111$.

Поскольку S является перестановкой, матрица C^{S} ортогональна. Действительно, евклидова норма каждого ее столбца равна 1, а любые два различных столбца ортогональны:

$$\sum_{w \in \mathbb{F}_2^3} C^{\mathsf{S}}_{w,u} C^{\mathsf{S}}_{w,v} = \delta^u(v).$$

Для всех $u, v \in \mathbb{F}_2^3$. То же самое справедливо для любых двух строк. Как и в доказательстве теоремы 2.4, $\delta^u : \mathbb{F}_2^n \to \mathbb{R}$ – функция, определенная как $\delta^u(u) = 1$ и $\delta^u(v) = 0$ для всех $v \neq u$. ▷

2.4. Корреляционные матрицы структурных функций

Как обсуждалось в разделе 1.1, криптографические функции являются композициями функций со специальной структурой, благодаря которой их можно вычислять эффективно. В этом разделе приводится корреляционная матрица для двух типов широкоупотребительных структурных функций.

К первому типу относятся линейные и, более общо, аффинные функции. Сложение с ключом и перестановка битов в демонстрационном шифре из раздела 1.1 – примеры аффинных функций.

Теорема 2.6. *Пусть* $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ *– аффинное отображение вида* $\mathsf{F}(x) = Ax + b$*, где* A *– матрица размера* $m \times n$ *над* $\mathbb{F}_2$*, а* b *– вектор, принадлежащий* $\mathbb{F}_2^m$*. Корреляционная матрица* C^{F} *отображения* F *удовлетворяет соотношению*

$$C^{\mathsf{F}}_{v,u} = (-1)^{v^{\mathsf{T}} b}\, \delta^u\left(A^{\mathsf{T}} v\right).$$

Доказательство. Доказательство проводится прямым вычислением:

$$C^{\mathsf{F}}_{v,u} = (-1)^{v^{\mathsf{T}} b}\, \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{(A^{\mathsf{T}} v + u)^{\mathsf{T}} x} = (-1)^{v^{\mathsf{T}} b}\, \delta^u\left(A^{\mathsf{T}} v\right).$$

Второе равенство следует из тождества $\sum_{x \in \mathbb{F}_2^n} (-1)^{w^{\mathsf{T}} x} = 2^n\, \delta^0(w)$, которое мы вывели в ходе доказательства теоремы 2.4. □

Теорему 2.6 можно интерпретировать следующим образом. Для заданной выходной маски у всякой линейной функции существует единственно возможная эффективная линейная аппроксимация, а ее входная маска является линейной функцией от заданной выходной маски. Кроме того, корреляция любой эффективной линейной аппроксимации равна 1. Заметим, что это верно, даже когда линейная функция не обратима. Для сложения с константой входная и выходная маски эффективной линейной аппроксимации должны быть равны, а корреляция равна ±1 в зависимости от значения константы.

Второй и последний класс структурных функций, обсуждаемый в этом разделе, иногда называют «кладочными (от выражения «каменная кладка») отображениями» (bricklayer map). На рис. 2.1 показана их структура.

Кладочное отображение – это функция $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$, построенная из l функций $\mathsf{F}_1, \ldots, \mathsf{F}_l$, которые применяются к непересекающимся частям входа. Точнее,

$$\mathsf{F}(x_1 \| x_2 \| \cdots \| x_l) = \mathsf{F}_1(x_1) \| \mathsf{F}_2(x_2) \| \cdots \| \mathsf{F}_l(x_l),$$

где $\|$ обозначает конкатенацию битовых векторов. Уровень S-блоков в демонстрационном шифре из примера 1.1 дает хороший пример.

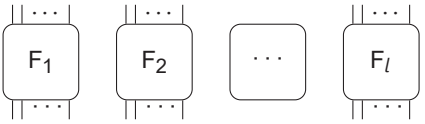


Рис. 2.1. Кладочная функция, построенная из l функций $\mathsf{F}_1, \ldots, \mathsf{F}_l$

Для описания структуры корреляционных матриц кладочных отображений полезно произведение Кронекера. Пусть A матрица размера $m \times n$ над полем $\mathbb{R}$:

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{bmatrix}.$$

Пусть также B – вещественная матрица размера $p \times q$. Произведением Кронекера матриц A и B называется блочная вещественная матрица $A \otimes B$ размера $pm \times qn$:

$$A \otimes B = \begin{bmatrix} A_{1,1}B & A_{1,2}B & \cdots & A_{1,n}B \\ A_{2,1}B & A_{2,2}B & \cdots & A_{2,n}B \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1}B & A_{m,2}B & \cdots & A_{m,n}B \end{bmatrix}.$$

Элементы $A \otimes B$ вычисляются по формуле:

$$(A \otimes B)_{p(i-1)+k,q(j-1)+l} = A_{i,j} B_{k,l}.$$

Альтернативно элементы $A \otimes B$ можно индексировать парами индексов. В этом случае имеем $(A \otimes B)_{(i,k),(j,l)} = A_{i,j} B_{k,l}$. Поскольку элементы корреляционных матриц индексируются битовыми векторами, используется следующее соглашение:

$$\left(C^{\mathsf{F}} \otimes C^{\mathsf{G}}\right)_{v_1\|v_2,u_1\|u_2} = C^{\mathsf{F}}_{v_1,u_1} C^{\mathsf{G}}_{v_2,u_2},$$

где C^{F} и C^{G} – корреляционные матрицы функций F и G соответственно.

Теорема 2.7. *Пусть* $\mathsf{F}_1, \ldots, \mathsf{F}_l$ *– функции* $\mathsf{F}_i : \mathbb{F}_2^{ni} \to \mathbb{F}_2^{mi}$, *и пусть* $n = \sum_{i=1}^{l} n_i$ *и* $m = \sum_{i=1}^{l} m_i$. *Корреляционная матрица кладочной функции* $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$, *определенной как* $\mathsf{F}(x_1\| \ldots \|x_l) = \mathsf{F}_1(x_1)\| \ldots \|\mathsf{F}_l(x_l)$, *равна*

$$C^{\mathsf{F}} = \bigotimes_{i=1}^{l} C^{\mathsf{F}_i}.$$

Доказательство. Обозначим $u = u_1\| \ldots \|u_l$ и $v = v_1\| \ldots \|v_l$. Элементы C^{F} равны

$$\begin{aligned}
C^{\mathsf{F}}_{v,u} &= \frac{1}{2^n} \sum_{x_1\|\cdots\|x_l \in \mathbb{F}_2^n} (-1)^{v_1^{\mathsf{T}}\mathsf{F}_1(x_1)+\cdots+v_l^{\mathsf{T}}\mathsf{F}_l(x_l)+u_1^{\mathsf{T}}x_1+\cdots+u_l^{\mathsf{T}}x_l} \\
&= \sum_{x_1 \in \mathbb{F}_2^{n_1}} \cdots \sum_{x_l \in \mathbb{F}_2^{n_l}} \prod_{i=1}^{l} \frac{1}{2^{n_i}} (-1)^{v_i^{\mathsf{T}}\mathsf{F}_i(x_i)+u_i^{\mathsf{T}}x_i} \\
&= \prod_{i=1}^{l} \frac{1}{2^{n_i}} \sum_{x \in \mathbb{F}_2^{n_i}} (-1)^{v_i^{\mathsf{T}}\mathsf{F}_i(x_i)+u_i^{\mathsf{T}}x_i} \\
&= \prod_{i=1}^{l} C^{\mathsf{F}}_{v_i,u_i}.
\end{aligned}$$

Теперь результат следует из определения произведения Кронекера. □

По-другому доказательство теоремы 2.7 можно рассматривать как применение леммы о набегании знаков в форме леммы 2.1. Это допустимо, потому что входы функций $\mathsf{F}_1, \ldots, \mathsf{F}_l$ независимы.

Пример 2.2 (сложение с ключом). Рассмотрим сложение с n-битовым ключом, т. е. функцию $x \mapsto x + k$ на $\mathbb{F}_2^n$. Для краткости корреляционную матрицу этой функции будем обозначать C^k. Эту операцию можно представлять себе как кладочное отображение, потому что i-й бит $x + k$ равен просто $x_i + k_i$. Поэтому из теоремы 2.7 следует, что

$$C^k = \bigotimes_{i=1}^{n} \begin{bmatrix} 1 & 0 \\ 0 & (-1)^{k_i} \end{bmatrix}.$$

Этот результат можно также получить с помощью теоремы 2.6. ▷

2.5. Линейные следы

Теорема 2.4 выражает корреляции всех линейных аппроксимаций композиции $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_1$ r функций в терминах корреляций линейных аппроксимаций r функций $\mathsf{F}_1, \ldots, \mathsf{F}_r$ по отдельности. В терминах корреляционных матриц имеем

$$C^{\mathsf{F}} = C^{\mathsf{F}_r} \cdots C^{\mathsf{F}_2} C^{\mathsf{F}_1}.$$

Хотя этот результат представляет теоретический интерес, практические вычисления трудны из-за большого размера корреляционных матриц. Чтобы обойти эту трудность, мы воспользуемся разреженностью корреляционных матриц. А именно запись приведенного выше произведения в терминах элементов приводит к следствию 2.8 ниже.

Следствие 2.8. *Пусть* $\mathsf{F}_1, \ldots, \mathsf{F}_r$ *– функции от битовых векторов. Корреляция линейной аппроксимации* $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_1$ *равна сумме корреляций всех линейных следов с такими же входной и выходной масками, как у аппроксимации:*

$$C^{\mathsf{F}}_{u_{r+1}, u_1} = \sum_{u_2, \ldots, u_r} \prod_{i=1}^{r} C^{\mathsf{F}_i}_{u_{i+1}, u_i}.$$

Если матрицы C^{F_i} разреженные, то сумма в следствии 2.8 содержит мало ненулевых членов. В общем случае идея в том, что ограниченное число следов определяет значение суммы с небольшой погрешностью. Это называется *аппроксимацией доминирующих следов*.

Традиционный принцип набегания знаков, который мы использовали в главе 1, корректен в предположении, что доминирует один след:

$$C^{\mathsf{F}}_{u_{r+1}, u_1} \approx \prod_{i=1}^{r} C^{\mathsf{F}_i}_{u_{i+1}, u_i},$$

где $(u_1, u_2, \ldots, u_{r+1})$ – след с наибольшей по абсолютной величине корреляцией. Это объясняет, почему лемма о набегании знаков иногда дает правильный результат, хотя лежащее в ее основе предположение о независимости не выполняется.

Для шифра с чередованием ключа $\mathsf{E}_k = \mathsf{R}_{k_r} \circ \cdots \circ \mathsf{R}_{k_1}$, где $\mathsf{R}_{k_i}(x) = \mathsf{R}(x) + k_i$, следствие 2.8 принимает вид

$$C^{\mathsf{E}_k}_{u_{r+1}, u_1} = \sum_{u_2, \ldots, u_r} (-1)^{\sum_{i=1}^{r} u_{i+1}^{\mathsf{T}} k_i} \prod_{i=1}^{r} C^{\mathsf{R}}_{u_{i+1}, u_i}. \tag{2.1}$$

В частности, раундовые ключи влияют на знаки корреляций следов, но не на их абсолютные величины.

Пример 2.3 (возврат к примеру 1.3). Как и в примере 1.3, рассмотрим линейную аппроксимацию (000000001, 000010000) трех раундов демонстрационного шифра из раздела 1.1. В примере 1.3 был найден линейный след с одинаковыми входной и выходной масками и корреляцией $(-1)^{k_0+k_{10}+k_{22}+k_{31}+1}/8$. В свете следствия 2.8 необходимо проверить, существуют ли другие следы, способные повлиять на корреляцию линейной аппроксимации.

Поскольку все битовые векторы $u \neq 001$ такие, что $C^{\mathsf{S}}_{u,001} \neq 0$, содержат по меньшей мере два ненулевых бита, как минимум два S-блока во втором раунде шифра должны иметь ненулевые входную и выходную маски, чтобы по-

лучить эффективный след. S-блоки с ненулевой выходной маской называются *активными*. Приглядевшись к корреляционной матрице уровня S-блоков внимательнее, мы увидим, что единственные возможности – $u \in \{101,011,111\}$. Каждая из них дает уникальный эффективный след. На рис. 2.2 показаны все три получающихся эффективных следа.

Корреляцию каждого следа можно вычислить, применив теоремы 2.6 и 2.7. Например, рассмотрим след на рис. 2.2а. Корреляция в первом раунде равна $(-1)^{k_0}C^{\mathsf{S}}_{101,001} = (-1)^{k_0+1}/2$. Для второго раунда из теоремы 2.7 следует, что корреляция равна

$$(-1)^{k_{10}+k_{16}} \underbrace{C^{\mathsf{S}}_{010,010}}_{-1/2} \underbrace{C^{\mathsf{S}}_{000,000}}_{1} \underbrace{C^{\mathsf{S}}_{010,010}}_{-1/2} = (-1)^{k_{10}+k_{16}}/4.$$

Наконец, в третьем раунде корреляция равна $(-1)^{k_{21}+k_{22}+k_{31}+1}/2$. Отсюда общая корреляция следа равна $(-1)^{k_0+k_{10}+k_{16}+k_21+k_{22}+k_{31}+1}/16$. Аналогичное вычисление для других следов дает полную корреляцию, равную (проверьте!)

$$(-1)^{\kappa_1}/8 + (-1)^{\kappa_1+\kappa_2}/16 + (-1)^{\kappa_1+\kappa_3}/16 + (-1)^{\kappa_1+\kappa_2+\kappa_3}/32,$$

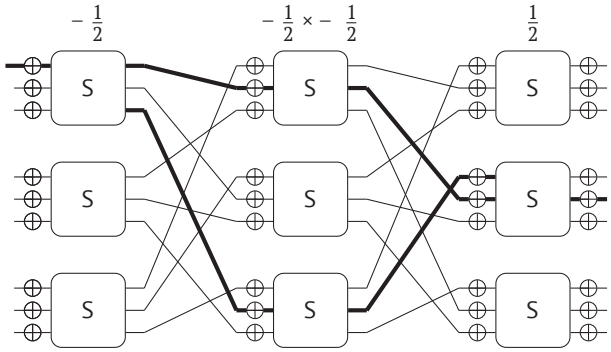
где $\kappa_1 = k_0 + k_{10} + k_{22} + k_{31} + 1$, $\kappa_2 = k_{16} + k_{21}$ и $\kappa_3 = k_{13} + k_{23}$. Поскольку других следов с одинаковыми входной и выходной масками нет, приведенное выше выражение точное. Заметим, что его можно переписать в виде

$$(-1)^{\kappa_1}/8\,\big(1 + (-1)^{\kappa_2}/2\big)\big(1 + (-1)^{\kappa_3}/2\big).$$

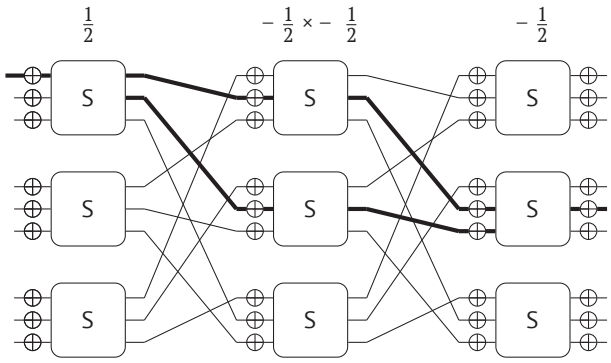
Исходя из вышесказанного, корреляция в зависимости от ключа равна либо ±1/32, либо ±3/32, либо ±9/32. В примере 1.4 было отмечено, что корреляция близка к 3/32. Это объясняется тем, что тогда использовался ключ, для которого $\kappa_1 = 0$, $\kappa_2 = 1$ и $\kappa_3 = 0$. ▷

2.6. Историческая справка

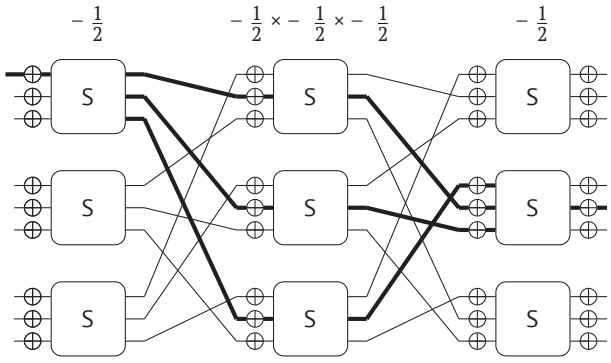
Корреляционные матрицы в 1994 году ввел Дамен на конференции по основаниям программной инженерии (FSE), а также в своей докторской диссертации, которую он защитил в следующем году. Важное отличие описания линейного криптоанализа, данного Мацуи, и на основе корреляционных матриц заключается в том, что Мацуи предполагал, что (раундовые) ключи – равномерно распределенные случайные величины. Уход от случайных ключей стал важным шагом вперед, и настоящая книга идет по тому же пути, но тогда он не был оценен по достоинству. В ранних работах по линейному криптоанализу часто рассматривались свойства квадратичных корреляций, усредненных по ключам. Самый важный пример такого результата – *теорема Нюберг о линейной оболочке* (см. упражнение 2.11).



(a) След с корреляцией $(-1)^{k_0+k_{10}+k_{16}+k_{21}+k_{22}+k_{31}+1}/16$



(b) След с корреляцией $(-1)^{k_0+k_{10}+k_{13}+k_{22}+k_{23}+k_{31}+1}/16$



(c) След с корреляцией $(-1)^{k_0+k_{10}+k_{13}+k_{21}+k_{22}+k_{23}+k_{31}+1}/32$

Рис. 2.2. Три следа с четырьмя или более активными S-блоками

2.7. Литература

Daemen, Joan (Mar. 1995). «Cipher and Hash Function Design Strategies Based on Linear and Differential Cryptanalysis». PhD thesis. KU Leuven.

Daemen, Joan, Rener Govaerts, and Joos Vandewalle (Dec. 1995). «Correlation Matrices». In: *FSE'94*. Ed. by Bart Preneel. Vol. 1008. LNCS. Springer, Berlin, Heidelberg, pp. 275–285. doi: 10.1007/3-540-60590-8_21.

Nyberg, Kaisa (May 1995). «Linear Approximation of Block Ciphers (Rump Session)». In: *EUROCRYPT'94*. Ed. by Alfredo De Santis. Vol. 950. LNCS. Springer, Berlin, Heidelberg, pp. 439–444. doi:10.1007/BFb0053460.

2.8. Упражнения

Упражнение 2.1

Пусть $f : \mathbb{F}_2^n \to \mathbb{F}_2$ и $g : \mathbb{F}_2^n \to \mathbb{F}_2$ – булевы функции. Докажите, что

$$\mathsf{C}(f,g) = \langle (-1)^f, (-1)^g \rangle.$$

Здесь $\langle \cdot,\cdot \rangle$ обозначает следующее скалярное произведение в векторном пространстве вещественных функций на $\mathbb{F}_2^n$:

$$\langle p, q \rangle = \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} p(x)\, q(x)$$

для $p : \mathbb{F}_2^n \to \mathbb{R}$ и $q : \mathbb{F}_2^n \to \mathbb{R}$.

Упражнение 2.2

Как показано в упражнении 1.2, любая линейная булева функция на $\mathbb{F}_2^n$ имеет вид ℓ_u, где $u \in \mathbb{F}_2^n$ и $\ell_u(x) = u^\mathsf{T} x$.

1. Докажите, что для всех $u, v \in \mathbb{F}_2^n$ $\mathsf{C}(\ell_u, \ell_v) = \delta^u(v)$.
2. Воспользовавшись этим свойством, перефразируйте доказательство теоремы 2.6 для линейных функций.

Упражнение 2.3

Пусть $f : \mathbb{F}_2^n \to \mathbb{F}_2$ – функция. Преобразованием Уолша–Адамара функции f называется функция $\mathcal{W}_f : \mathbb{F}_2^n \to \mathbb{R}$, определенная следующим образом:

$$\mathcal{W}_f(u) = \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{u^\mathsf{T} x + f(x)}.$$

Докажите следующие утверждения:

1. $\mathcal{W}_f(u) = C(\ell_u, f)$.
2. $\mathcal{W}_f(u) = 0$ тогда и только тогда, когда $f + \ell_u$ является сбалансированной булевой функцией. Булева функция называется сбалансированной, если она принимает значение 0 ровно на половине своих входов.
3. Корреляция линейной аппроксимации (u, v) функции $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ равна $\mathcal{W}_{\ell_v \circ \mathsf{F}}(u)$.

Упражнение 2.4

В этом упражнении исследуется ряд интересных следствий теорем 2.4–2.7. Все приведенные ниже результаты вытекают из одной из этих теорем. Пусть $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^n$ – перестановка. Докажите следующие утверждения.

1. Для любой $u \in \mathbb{F}_2^n$ преобразование Уолша–Адамара (см. упражнение 2.3) удовлетворяет тождествам $\sum_{v\in\mathbb{F}_2^n} \mathcal{W}_{\ell_u\circ \mathsf{F}}(v)^2 = 1$ и $\sum_{v\in\mathbb{F}_2^n} \mathcal{W}_{\ell_u\circ \mathsf{F}}(u)^2 = 1$. В частности, из первого утверждения следует, что $\sum_{v\in\mathbb{F}_2^n} \mathcal{W}_f(v)^2 = 1$ для любой булевой функции f. Иногда этот факт называют равенством Парсеваля.
2. Для любой ненулевой $v \in \mathbb{F}_2^n$ булева функция $\ell_v \circ \mathsf{F}$ сбалансирована.

* Упражнение 2.5

1. Придумайте алгоритм с временной сложностью $O_s(l\, s^l)$, который вычисляет произведение вектора на матрицу B вида

$$B = A_1 \otimes \ldots \otimes A_l,$$

где $A_1, \ldots, A_l$ – матрицы размера $s\times s$ и $l \geqslant 1$ – целое число.

2. Докажите, что для любой функции $f : \mathbb{F}_2^n \to \mathbb{F}_2$ (о том, что такое $\mathcal{W}_f$, см. упражнение 2.3)

$$\begin{bmatrix} \mathcal{W}_f(0,0,\ldots,0) \\ \mathcal{W}_f(0,0,\ldots,1) \\ \vdots \\ \mathcal{W}_f(1,1,\ldots,1) \end{bmatrix} = \frac{1}{2^n}\left(\bigotimes_{i=1}^{n} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right) \begin{bmatrix} (-1)^{f(0,0,\ldots,0)} \\ (-1)^{f(0,0,\ldots,1)} \\ \vdots \\ (-1)^{f(1,1,\ldots,1)} \end{bmatrix}.$$

3. Придумайте алгоритм, вычисляющий $\mathcal{W}_f$ за время $O(n2^n)$.
4. Придумайте алгоритм, вычисляющий корреляционную матрицу заданной функции $\mathsf{F}\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ за время $O(n2^{n+m})$ в предположении, что F задана в табличной форме.

Упражнение 2.6

В этом упражнении вам предлагается проанализировать конструкцию на рис. 2.3. Ее вход обозначается x, а секретный ключ – k. Корреляционная матрица S-блока S показана в примере 2.1.

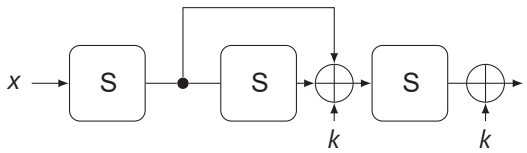


Рис. 2.3. Конструкция с тремя S-блоками

1. Найдите линейный след с корреляцией $\pm 1/4$.
2. Найдите линейную аппроксимацию с корреляцией 1 хотя бы для одного ключа.
3. Предположим, что существует вход x, которому соответствует выход 001. Исходя из вашего ответа на предыдущий вопрос, скажите, каковы возможные значения ключа.

Упражнение 2.7

В этом упражнении вам предлагается проанализировать конструкцию $\mathsf{E}_k : \mathbb{F}_2^6 \to \mathbb{F}_2^6$ на рис. 2.4. Ее секретный ключ обозначен $k = k_1 \| k_2$. Корреляционная матрица S-блока S приведена в примере 2.1.

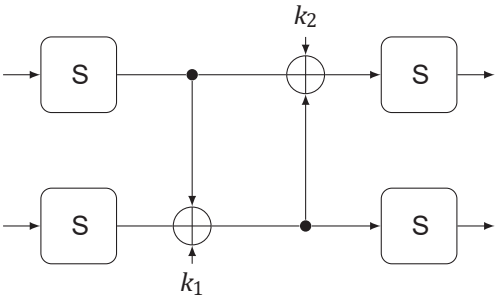


Рис. 2.4. Конструкция с четырьмя S-блоками

1. Найдите линейный след с корреляцией $\pm 1/4$.
2. Найдите нетривиальную линейную аппроксимацию с корреляцией 1 для $k = 000000$. Вычислите ее корреляцию для всех значений k.
3. Исходя из вашего ответа на предыдущий вопрос, скажите, возможно ли, что $k_1 = k_2$?

Упражнение 2.8

Обозначим $\mathrm{and}_n : \mathbb{F}_2^{2n} \to \mathbb{F}_2^n$ функцию, вычисляющую поразрядное И своих n-битовых входов:

$$\mathrm{and}_n : (x_1, \ldots, x_n, y_1, \ldots, y_n) \mapsto (x_1 y_1, \ldots, x_n y_n).$$

Эта операция используется в некоторых блочных шифрах (см. упражнение 2.9). Следующие вопросы ведут к выводу формулы элементов корреляционной матрицы and_n.

1. Покажите, что $1 + (-1)^a + (-1)^b - (-1)^{a+b} = 2(-1)^{ab}$ для любых $a, b \in \mathbb{F}_2$.
2. Вручную вычислите корреляционную матрицу and_1.
3. Покажите, что для любого $n \geqslant 1$ элементы корреляционной матрицы and_n описываются формулой

$$C^{\mathrm{and}_n}_{w,u\|v} = \begin{cases} (-1)^{u^{\mathsf{T}} v}/2^{\mathrm{wt}(w)}, & \text{если } u \preccurlyeq w \text{ и } v \preccurlyeq w, \\ 0 & \text{в противном случае,} \end{cases}$$

где $\mathrm{wt}(w)$ – количество ненулевых элементов w (его часто называют «весом Хэмминга» w), а $\preccurlyeq$ обозначает порядок на битовых строках, определенный как $x \preccurlyeq y$ тогда и только тогда, когда $x_i \leqslant y_i$ для всех $i = 1, \ldots, n$.

* Упражнение 2.9

Simon – так называемый шифр Фейстеля, спроектированный в Агентстве национальной безопасности (АНБ) США. Он основан на функции поразрядного И из

упражнения 2.8. Раундовая функция шифра Simon-32 показана на рис. 2.5. Поразрядное И обозначается символом ∧, а поразрядный циклический сдвиг – символом ≪. В этом упражнении нам не важно, как генерируются раундовые ключи k_i.

1. Найдите линейный след в двух раундах с максимальной абсолютной корреляцией.
2. Реализуйте шифр Simon-32 и оцените корреляцию линейной аппроксимации, найденной вами в предыдущем вопросе, для ключа, состоящего из одних нулей. Согласуются ли результаты с вашими ожиданиями? Попробуйте объяснить потенциальные расхождения.

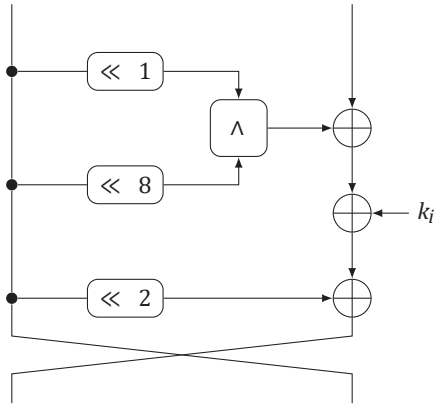


Рис. 2.5. Раундовая функция блочного шифра Simon

* Упражнение 2.10

Функция $f : \mathbb{F}_2^n \to \mathbb{F}_2$ называется *квадратичной формой*, если она имеет вид

$$f(x) = \sum_{i<j} a_{i,j}\, x_i x_j,$$

где множители $a_{i,j}$ принадлежат $\mathbb{F}_2$ и $x_1, \ldots, x_n$ – элементы x.

1. Докажите, что существует верхнетреугольная матрица A такая, что $f(x) = x^{\mathsf{T}} A x$.
2. Покажите, что с точностью до линейной замены переменных матрица A имеет вид:

$$\begin{bmatrix} 0 & 1 & & & & & & & \\ 0 & 0 & & & & & & & \\ & & 0 & 1 & & & & & \\ & & 0 & 0 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 0 & 1 & & \\ & & & & & 0 & 0 & & \\ & & & & & & & 0 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 0 \end{bmatrix}.$$

3. Воспользовавшись теоремой 2.6 и результатом упражнения 2.8, выведите формулу для корреляционной матрицы произвольной квадратичной формы f (это матрица размера 2×2^n). Докажите, что абсолютная величина корреляции любой нетривиальной линейной аппроксимации f не превосходит 2^{-r}, где r – ранг A.
4. Пусть $\mathsf{F}\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ определена как $\mathsf{F}(x) = \mathsf{Q}(x) + \mathsf{L}(x)$, где функция L линейная, а все элементы Q являются квадратичными формами. Придумайте алгоритм, вычисляющий корреляцию заданной линейной аппроксимации F за время $O(n^3)$.

Упражнение 2.11

Рассмотрим шифр с чередованием ключа $\mathsf{E}_{\mathbf{k}}$ с раундовой функцией R и независимыми и равномерно распределенными раундовыми ключами $(\mathbf{k}_1, \ldots, \mathbf{k}_r) = \mathbf{k}$.. Докажите следующие утверждения.

1. $\mathbb{E}_{\mathbf{k}}(C^{\mathsf{E}_{\mathbf{k}}}_{v,u}) = 0$ для любых $(u, v) \neq (0, 0)$.
2. Для любых (u_{r+1}, u_1) дисперсия равна

$$\mathbb{E}_{\mathbf{k}}\left(C^{\mathsf{E}_{\mathbf{k}}}_{u_{r+1},u_1}\right)^2 = \sum_{u_2,\ldots,u_r} \prod_{i=1}^{r} \left(C^{\mathsf{R}}_{u_{i+1},u_i}\right)^2.$$

Этот результат называется *теоремой Нюберг о линейной оболочке*.

Глава 3

Оптимизация линейных следов

Нахождение линейных следов с высокой абсолютной корреляцией быстро становится утомительным занятием, особенно для шифров с более сложной структурой, чем пример, с которым мы работали до сих пор. Поскольку общее число следов конечно, нахождение линейных следов с максимальной абсолютной корреляцией – пример задачи комбинаторной оптимизации.

В этой главе обсуждаются три распространенных метода оптимизации: метод ветвей и границ Мацуи, смешанно-целочисленное линейное программирование и выполнимость или выполнимость формул в теориях. Попутно вводятся два дополнительных примера шифров с разными стратегиями проектирования.

3.1. Метод ветвей и границ

Первый метод оптимизации, который мы обсудим, принадлежит Мацуи. Это пример алгоритма ветвей и границ при поиске в глубину, но можно предложить многочисленные варианты базовой стратегии, дающие преимущество в конкретных случаях.

3.1.1. Поиск в глубину

Алгоритм поиска в глубину обходит вершины графа, следуя по ребрам настолько далеко, насколько удастся, а затем выполняя возврат. На рис. 3.1 этот процесс показан для случая, когда граф является деревом. Начав с корня a, алгоритм на каждом шаге выбирает одного из сыновей ранее выбранной вершины. На рис. 3.1 одна за другой посещаются вершины a, b, c и d. У вершины d нет потомков, поэтому алгоритм возвращается к последней вершине, имеющей неисследованных потомков. Это вершина b, из которой поиск продолжается посещением вершин e и f. Поскольку у f нет потомков, алгоритм возвращается к вершине e. После посещения g он возвращается в корень и продолжает исследовать вершины h, i и j.

Нахождение кратчайших путей в реберно взвешенном графе – типичное приложение алгоритмов поиска в глубину. Кратчайшим путем между двумя вершинами называется последовательность соединенных ребрами вершин такая, что сумма весов этих ребер минимальна. Можно рассмотреть более общую задачу о поиске кратчайших путей между двумя множествами вершин. Например, обход в глубину, показанный на рис. 3.1, можно было бы использовать для нахождения кратчайшего пути между корнем a и одной из листовых

вершин *d*, *f*, *g* или *j*, для чего нужно было бы сохранять длину пути, по которому следует алгоритм. Чтобы найти кратчайший путь, этот метод посещает каждую вершину как минимум один раз. Однако можно надеяться найти достаточно короткий путь раньше, применяя эвристику для выбора следующей подлежащей посещению вершины. Одна из возможных стратегий – всегда следовать по ребру с наименьшим весом. Она называется жадным поиском.

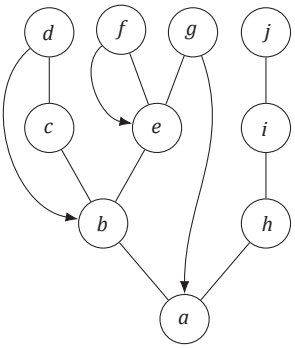


Рис. 3.1. Обход дерева высоты 3 в глубину

Метод ветвей и границ находит кратчайший путь, не посещая все вершины. Он основан на предположении о том, что продолжение пути никогда не уменьшает его полный вес. Если все веса неотрицательны, то это предположение заведомо верно. Метод ветвей и границ можно объединить с поиском в глубину, модифицировав выбор следующей вершины. Пусть x – текущая вершина, $f(x)$ – полный вес текущего пути от корня к x, а B – полный вес наилучшего пути, найденного к настоящему моменту. В алгоритме также используется нижняя граница $h(y)$ полного веса кратчайшего пути, начинающегося в вершине y. Функция $h\colon V \to \mathbb{R}$ называется эвристикой. Сына y вершины x можно игнорировать, если

$$f(x) + w_{x,y} + h(y) \geqslant B,$$

где $w_{x,y}$ – вес ребра, соединяющего x и y. Действительно, $f(x) + w_{x,y} + h(y)$ – нижняя граница полного веса всех путей, которые могли бы быть найдены в результате продолжения поиска из вершины y.

Количество вершин, посещенных методом ветвей и границ с поиском в глубину, сильно зависит от эвристической функции h. В худшем случае $h(y) = 0$ для всех вершин y.

3.1.2. Метод Мацуи

Пусть $\mathsf{F} = \mathsf{F}_r \circ \ldots \circ \mathsf{F}_1$ – композиция r функций $\mathsf{F}_1, \ldots, \mathsf{F}_r$, где $\mathsf{F}_i : \mathbb{F}_2^{n_i} \to \mathbb{F}_2^{n_{i+1}}$. Напомним, что следом называется последовательность $(u_1, \ldots, u_{r+1})$ масок. Множество всех возможных следов образует граф G с множеством вершин

$$V = \bigcup_{1 \leq i \leq r+1} \left\{ (i, u_i) \mid u_i \in \mathbb{F}_2^{n_i} \right\}.$$

Ребрами соединены вершины (i, u_i) и $(i + 1, u_{i+1})$, для которых $C^{\mathsf{F}_i}_{u_{i+1},u_i} \neq 0$, а вес ребра равен $-\log_2 |C^{\mathsf{F}_i}_{u_{i+1},u_i}|$. Если ребро ориентировано (направлено из $i + 1$ в i), то G является ориентированным ациклическим графом. Полный вес пути $(1, u_1), (2, u_2), \ldots, (r + 1, u_{r+1})$ равен

$$\sum_{i=1}^{r} -\log_2 \left|C^{\mathsf{F}_i}_{u_{i+1},u_i}\right| = -\log_2 \prod_{i=1}^{r} \left|C^{\mathsf{F}_i}_{u_{i+1},u_i}\right|.$$

Следовательно, кратчайший путь между вершинами $(1, u_1)$ и $(r + 1, u_{r+1})$ соответствует следу с максимальной абсолютной корреляцией. Метод Мацуи находит такие следы с помощью поиска в глубину, начиная с последнего раунда и двигаясь к первому раунду[1]. Псевдокод метода приведен в алгоритме 3.1.

Эвристическая функция h: $V \to \mathbb{R}$, используемая в алгоритме 3.1, определяется в терминах границ $B_1, \ldots, B_r$ максимальной абсолютной корреляции следа из раунда 1 в раунд i. То есть

$$h(i, u_i) = -\log_2 B_i.$$

На практике эвристическую функцию часто можно улучшить, приняв во внимание конкретное значение маски u_i. Заметим, что всегда можно выбрать $B_i = 1$.

Порядок перечисления масок (строки 10 и 17) в алгоритме 3.1 оставлен неопределенным, но на практике он важен. Во внутреннем цикле (строка 10) обычно используется жадный подход: маски u_{l-1} выбираются в порядке убывания $|C^{\mathsf{F}_i}_{u_{i+1},u_i}|$. Однако при разрешении неоднозначностей возможны варианты. Во внешнем цикле (строка 17) необходим другой подход. Обычно используется эвристика, основанная на заглядывании вперед на один раунд. Например, можно посчитать количество S-блоков, которые гарантированно активны в последнем раунде.

Алгоритм 3.1. Псевдокод метода Мацуи нахождения следов

Вход:

Функции $\mathsf{F}_1, \ldots, \mathsf{F}_r$ с корреляционными матрицами $C^{\mathsf{F}_1}, \ldots, C^{\mathsf{F}_r}$

Границы $B_0 = 1, B_1, \ldots, B_r$ такие, что $B_l \geqslant B_l \geq \prod_{i=1}^{l-1} \left|C^{\mathsf{F}_i}_{u_{i+1},u_i}\right|$ для всех $u_1, \ldots, u_l$

Выход:

След $(v_1, \ldots, v_{r+1})$ с максимальной величиной $\prod_{i=1}^{r} \left|C^{\mathsf{F}_i}_{v_{i+1},v_i}\right|$

1 ▷ Инициализировать абсолютную корреляцию лучшего из найденных к настоящему моменту следов
2: $B \leftarrow 0$
3: ▷ Рекурсивная процедура нахождения лучшего следа, начиная с последнего раунда
4: **procedure** SEARCH($u_{r+1}, u_r, \ldots, u_l$)

---

[1] Оригинальный метод Мацуи начинает с первого раунда, но работать в обратном направлении лучше, когда некоторые из функций $\mathsf{F}_1, \ldots, \mathsf{F}_r$ не обратимы.

```
5       if l = 1 then
6:          B ← ∏_{i=1}^{r} |C^{F_i}_{u_{i+1},u_i}|
7:          (v_1, ..., v_{r+1}) ← (u_1, ..., u_{r+1})
8:          return
9:      end if
10:     for всех u_{l-1} таких, что C^{F_i}_{u_{i+1},u_i} ≠ 0 do      ▷ Эвристика определения порядка
11:         if B_{l-1} ∏_{i=l-1}^{r} |C^{F_i}_{u_{i+1},u_i}| > B then
12:             SEARCH (u_{r+1}, u_r, ... , u_l, u_{l-1})
13:         end if
14:     end for
15: end procedure
16:
17: for всех u_{r+1} ≠ 0 do                                          ▷ Эвристика определения порядка
18:     SEARCH (u_{r+1})
19: end for
20:
21: return (v_1, ..., v_{r+1})
```

Пример 3.1. Применим алгоритм 3.1 к трем раундам демонстрационного шифра из раздела 1.1. Выберем $B_1 = 1$. Абсолютная корреляция аппроксимации на одном раунде не превосходит $1/2$, потому что хотя бы один S-блок должен быть активен. Поэтому $B_2 = 1/2$ и $B_3 = 1/4$ – допустимые варианты выбора.

Обозначим u_1, u_2 и u_3 маски на входе в первые три раунда шифра соответственно. Выберем маску u_4, так чтобы количество S-блоков в последнем раунде было минимально, неоднозначность разрешаем случайным образом. Например, так можно было бы прийти к выбору $u_4 = 000010000$. При подобном выборе значениями-кандидатами u_3 будут {000010000, 000011000, 000101000, 000111000}. Снова разрешая неоднозначность случайным образом, предположим, что выбрано 000011000. Тогда возможными значениями u_2 будут

$$\{010000010, 011000010, 101000010, 111000010, 010000011, \ldots\}.$$

Снова разрешая неоднозначность случайным образом, выбираем значение 011000010. В этот момент все допустимые значения u_1 приводят к следу с абсолютной корреляцией $1/16$. Найдя такой след, алгоритм возвращается к выбору u_2. Однако так как $B_2/8 = 1/16$, никакой другой выбор u_2 не может дать лучший след. В результате алгоритм сразу возвращается к выбору u_3. Поскольку $B_3/2 = 1/8$, все остальные возможные значения u_3 по-прежнему являются допустимыми кандидатами. Однако для любого выбора, кроме 000010000, алгоритм немедленно выполняет возврат.

После того как в качестве значения u_3 выбрано 000010000, единственный выбор u_2, который не был признан тупиковым и исключен, – 000000010. Тогда любой допустимый выбор u_1 дает след с корреляцией $1/8$, который является оптимальным. Алгоритм немедленно выполняет возврат к выбору u_4. Рано прерывая поиск на основе количества S-блоков, активных в последнем раунде, алгоритм может завершиться после проверки еще $3 \cdot 8 - 2 = 22$ значений u_4. ▷

Метод ветвей и границ Мацуи применим также к другим типам шифров, включая примеры, которые будут приведены в разделах 3.2.1 и 3.3.1. Однако это утомительная и чреватая ошибками работа, особенно когда на первый план выходит эффективность. Альтернативный подход – переформулировать задачу оптимизации как задачу смешанно-целочисленного линейного программирования (раздел 3.2) или задачу выполнимости (раздел 3.3), чтобы их можно было решить, воспользовавшись готовыми программами.

3.2. Смешанно-целочисленное линейное программирование

Задача линейного программирования – это задача оптимизации в вещественных переменных $x_1, \ldots, x_n$ с линейной целевой функцией

$$\min_{x_1,\ldots,x_n} \sum_{i=1}^{n} a_i x_i,$$

где $a_1, \ldots, a_n$ – заданные вещественные числа. На переменные $x_1, \ldots, x_n$ накладывается произвольное число ограничений в виде линейных неравенств вида

$$\sum_{i=1}^{n} b_i x_i \geq b_{n+1},$$

где $b_1, \ldots, b_n$ и b_{n+1} – вещественные числа. Задачи линейного программирования допускают практически эффективное решение, например с помощью симплекс-метода. Если требуется, чтобы некоторые из переменных $x_1, \ldots, x_n$ были целыми, то говорят о задаче смешанно-целочисленного линейного программирования. Для приложений в области линейного криптоанализа целые переменные обычно принимают значения 0 или 1. В общем случае задачи линейного программирования с целыми переменными гораздо труднее. Тем не менее существуют специализированные программы, способные справиться с такими задачами на практике. В этом разделе мы примем существование таких «решателей» как данность, а сами займемся тем, как сформулировать задачу оптимизации следов в виде задачи смешанно-целочисленного линейного программирования.

3.2.1. Пример: шифр типа Rijndael

В этом разделе описывается второй демонстрационный шифр, который мы проанализируем методом смешанно-целочисленного линейного программирования в разделах 3.2.2 и 3.2.3. Это шифр с чередованием ключа, при проектировании которого применена *стратегия широкого следа*, чтобы противостоять линейному криптоанализу.

Общая структура раундовой функции показана на рис. 3.2. Это функция на $\mathbb{F}_2^{96}$, но описывать раундовую функцию удобнее, когда биты состояния развернуты на сетке 4×8 в 3-битовые группы («ячейки»). Первая операция – кладочная функция, или уровень S-блоков, она называется `SubCells`. Второй и третий шаги – линейные функции: на втором шаге, названном `ShiftRows`, строки состояния циклически сдвигаются, а на третьем – `MixColumns` – линейная функция применяется к каждому столбцу. Таким образом, бесключевая раундовая функция имеет вид

$\mathsf{R} = \texttt{MixColumns} \circ \texttt{ShiftRows} \circ \texttt{SubCells}.$

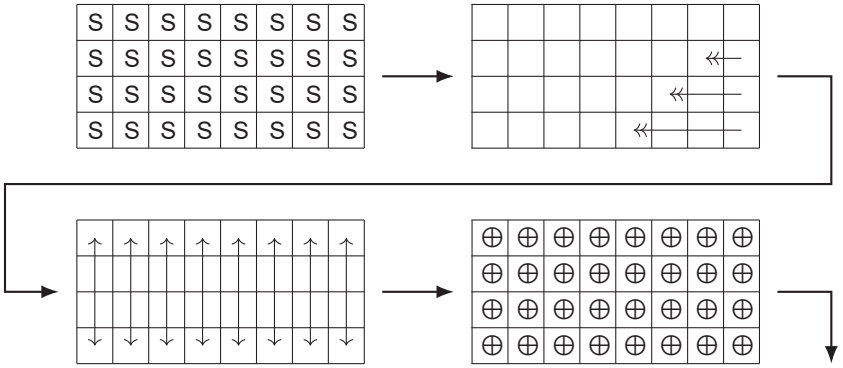


Рис. 3.2. Раундовая функция состоит из функций SubCells, ShiftRows, MixColumns и сложения с раундовым ключом. Источник: Beierle et al. (2018). © IACR

Ниже приведены дополнительные сведения о функциях в правой части.

SubCells заключается в параллельном применении S-блока S к 3-битовым ячейкам состояния. S-блок S: $\mathbb{F}_2^3 \to \mathbb{F}_2^3$ такой же, как в демонстрационном шифре из раздела 1.1.

ShiftRows переставляет ячейки состояния. Если строки пронумерованы от нуля до трех и 0 соответствует верхней строке, то `ShiftRows` циклически сдвигает i-ю строку состояния на $3 \cdot i$ бит влево.

MixColumns применяет линейное отображение к каждому столбцу состояния. Обозначим $(x_1, \ldots, x_4)$, где $x_i \in \mathbb{F}_2^3$, столбец состояния. `MixColumns` отображает $(x_1, \ldots, x_4)$ в новый столбец следующим образом (здесь I – единичная матрица размера 3×3):

$$\begin{bmatrix} 0 & I & I & I \\ I & 0 & I & I \\ I & I & 0 & I \\ I & I & I & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

В упражнении 3.1 вам будет предложено доказать, что это отображение обратимо.

i-я раундовая функция определяется как $\mathsf{R}_{k_i}(x) = \mathsf{R}(x) + k_i$, а шифр является композицией $\mathsf{R}_{k_1}, \ldots, \mathsf{R}_{k_r}$, которой предшествует начальное сложение с раундовым ключом k_0. Пока что вопрос о количестве раундов r оставим открытым. Раундовые ключи имеют вид $k_i = k + c_i$, $i = 0, \ldots, r$, где $k \in \mathbb{F}_2^{96}$ – ключ шифра, а $c_1, \ldots, c_r$ – константы. Значения констант для ячеек состояния (слева направо и сверху вниз, с первого до последнего раунда) определяются последовательностью Туэ–Морса 000, 111, 111, 000, 111, 000, 000, 111 ….

Линейный уровень и шаг `MixColumns`, в частности, выбирается так, чтобы каждый линейный след гарантированно содержал много активных S-блоков. В этом состоит важное отличие от демонстрационного шифра из раздела 1.1,

для которого существуют линейные следы, содержащие всего один активный S-блок в раунде. По теореме 2.6, все эффективные линейные аппроксимации линейной функции $x \mapsto Mx$, где $M \in \mathbb{F}_2^{n \times n}$, имеют вид $(M^{\intercal}u, u)$, где $u \in \mathbb{F}_2^n$. Поэтому имеет смысл выбирать матрицу M так, чтобы общее число ненулевых ячеек в u и $M^{\intercal}u$ было большим. Действительно, эти ненулевые ячейки соответствуют активным S-блокам в раундах до и после линейного уровня. Отсюда вытекает следующее определение.

Определение 3.1 (число линейных разветвлений). Пусть M – матрица размера $bm \times bn$ над $\mathbb{F}_2$. Числом линейных разветвлений M (или функции $x \mapsto Mx$) называется величина

$$\mathcal{B}(M) = \min_{x \neq 0} \mathrm{wt}_m(x) + \mathrm{wt}_n(M^{\intercal}x),$$

где $\mathrm{wt}_l(x_1 \| x_2 \| \ldots \| x_l) = |\{1 \leqslant i \leqslant l \mid x_i \neq 0\}|$ – вес Хэмминга.

В упражнении 3.1 вам будет предложено показать, что число разветвлений функции `MixColumns` равно четырем. Это означает, что линейный след двух раундов всегда включает по меньшей мере четыре активных S-блока. В упражнении 3.3 вам будет предложено доказать, что любой четырехраундовый след включает по меньшей мере 16 активных S-блоков.

3.2.2. Построение модели

Чтобы сформулировать задачу оптимизации как смешанно-целочисленную линейную программу, мы введем двоичные целые переменные, соответствующие битам входной и выходной масок каждого уровня S-блоков. Моделирование шага `ShiftRows` не вызывает сложностей, поскольку сводится к перестановке (переименованию) переменных. Моделирование `MixColumns` и `SubCells` обсуждается ниже.

Шаг `MixColumns` основан на операциях ИСКЛЮЧАЮЩЕЕ ИЛИ. Условие, согласно которому z является результатом применения ИСКЛЮЧАЮЩЕГО ИЛИ к x и y, можно выразить с помощью линейных неравенств несколькими способами. Например, в предположении, что x, y и z – двоичные переменные,

$$\begin{aligned} x + y + z &\leqslant 2, \\ x + y + z &\geqslant 2d, \\ d &\geqslant x, \\ d &\geqslant y, \\ d &\geqslant z, \end{aligned}$$

где все сложения производятся в целых числах, а d – новая целая фиктивная переменная. Если d – двоичная переменная, то годится также следующее линейное равенство:

$$x + y + z = 2d.$$

Для реализации M нужно ИСКЛЮЧАЮЩЕЕ ИЛИ трех переменных. Чтобы выразить w в виде ИСКЛЮЧАЮЩЕГО ИЛИ x, y и z, это равенство можно обобщить:

$$w + x + y + z = 4d_1 - 2d_2 - 2d_3,$$

где d_1, d_2 и d_3 – фиктивные двоичные переменные. В частности, если (v_1, v_2, v_3, v_4) – выходная маска для $x \mapsto Mx$ и (u_1, u_2, u_3, u_4) – входная маска, то

$$\begin{aligned} u_{1,i} + v_{2,i} + v_{3,i} + v_{4,i} &= 4d_1 - 2d_2 - 2d_3, \\ u_{2,i} + v_{1,i} + v_{3,i} + v_{4,i} &= 4e_1 - 2e_2 - 2e_3, \\ u_{3,i} + v_{1,i} + v_{2,i} + v_{4,i} &= 4f_1 - 2f_2 - 2f_3, \\ u_{4,i} + v_{1,i} + v_{2,i} + v_{3,i} &= 4g_1 - 2g_2 - 2g_3, \end{aligned}$$

где $u_{1,i}$ – i-й бит u_1, и аналогично для u_2, u_3, u_4 и $v_1, \ldots, v_4$.

Для моделирования `SubCells` введем для каждого S-блока новую переменную, показывающую, является ли он активным. Кроме того, следует исключить неэффективные линейные аппроксимации S-блока. Предположим, что (u_1, u_2, u_3) – входная маска, а (v_1, v_2, v_3) – выходная маска. Если a – двоичная переменная, равная 1, когда S-блок активен, и 0 в противном случае, то

$$3a \geqslant v_1 + v_2 + v_3,$$

где под сложением снова понимается сложение целых чисел. Исключить линейные аппроксимации S-блока с нулевой корреляцией труднее. Однако существует общий подход к этой задаче. Напомним, что множество $S \subset \mathbb{R}^n$ называется выпуклым, если для любых $x, y \in S$ все точки, лежащие на соединяющем их отрезке $\{\lambda x + (1-\lambda)y \mid \lambda \in [0,1]\}$, также принадлежат S. Выпуклой оболочкой множества $T \subseteq \mathbb{R}^n$ называется наименьшее выпуклое множество, содержащее T.

Выпуклой оболочкой конечного множества является выпуклый политоп. Например, выпуклая оболочка множества $\{(0,0), (1,0), (0,1)\} \subset \mathbb{R}^2$ показана на рис. 3.3. Выпуклые политопы обладают полезным свойством – их можно описать конечным числом линейных неравенств. Это следует из того, что каждое линейное неравенство описывает полуплоскость.

Поэтому множество неравенств, исключающее неэффективные линейные аппроксимации S, можно найти, отыскав линейные неравенства, которые описывают выпуклую оболочку этого множества (с точностью до канонического отображения $\mathbb{F}_2 \hookrightarrow \{0,1\} \subset \mathbb{R}$)

$$T = \left\{ (u_1, u_2, u_3, v_1, v_2, v_3) \in \mathbb{F}_2^6 \;\middle|\; C^{\mathsf{S}}_{v_1 \| v_2 \| v_3, u_1 \| u_2 \| u_3} \neq 0 \right\} \subset \mathbb{R}^6.$$

Важно, что выпуклая оболочка T не содержит никаких точек $\{0,1\}^6$, кроме принадлежащих T. Не вдаваясь в детали, скажем, что выпуклую оболочку множества можно вычислить с помощью *алгоритма Quickhull*. Множество линейных неравенств, полученных из представления выпуклой оболочки, не обязательно минимально. Поэтому часто применяются дополнительные методы, чтобы уменьшить число линейных неравенств. Нахождение минимального подмножества неравенств – задача о покрытии множества, и ее саму можно решить методами смешанно-целочисленного линейного программирования.

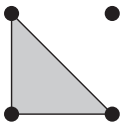


Рис. 3.3. Выпуклая оболочка множества {(0,0), (1,0), (0,1)}

Если $a_1, \ldots, a_l$ – двоичные переменные, показывающие, активен S-блок или нет, то целевая функция имеет вид

$$\sum_{i=1}^{l} a_i.$$

Она равна весу линейного следа, выраженному с помощью всех переменных. При этом предполагается, что ненулевые корреляции всех линейных аппроксимаций S-блока равны $\pm 1/2$. В общем случае может оказаться необходимым закодировать веса целыми переменными или определить несколько переменных, соответствующих линейным аппроксимациям с разными весами.

3.2.3. Решение модели

К популярным решателям задачи смешанно-целочисленного программирования (MILP) относятся CPLEX и Gurobi. Задачи подаются на вход решателя либо программно с помощью API, либо в виде файла. Большинство решателей поддерживают формат LP. LP-файл является текстовым файлом, содержащим целевую функцию, неравенства и типы всех переменных. Например:

```
\ Целевая функция (minimize или maximize)
minimize
x + y + z
\ Список неравенств
subject to
2 - x - y - z >= 0
x + y + z - 2d >= 0
d - x >= 0
d - y >= 0
d - z >= 0
\ Типы переменных
generals
d
binary
x y z
end
```

3.3. Выполнимость и выполнимость в теориях

Выполнимость (satisfiability), или «SAT», – это задача принятия решений, в которой спрашивается, можно ли назначить переменным значения «ложь» или «истина», так чтобы заданная булева формула стала истинной. На практике часто полезен вариант данной задачи, подразумевающий поиск, но правильное назначение можно эффективно найти, повторно решая задачу принятия решений. Большинство решателей предполагают, что булева формула является *конъюнктивной нормальной формой*. А именно любую булеву формулу с n переменными $x_1, \ldots, x_n$ над $\mathbb{F}_2$ можно записать в виде

$$\bigwedge_{i=1}^{m} (x_{i_1} + b_{i_1}) \vee (x_{i_2} + b_{i_2}) \vee \cdots \vee (x_{i_{l_i}} + b_{i_{l_i}}),$$

где $\wedge$ обозначает И, $\vee$ обозначает ИЛИ, а $b_{i1}, \ldots, bil_i \in \mathbb{F}_2$ – константы.

Выполнимость в теориях (satisfiability modulo theories – SMT) обобщает задачу выполнимости на более общие формулы, которые могут включать кванторы, целые переменные, битовые векторы и т. д. На внутреннем уровне SMT-решатели часто преобразуют по крайней мере часть задачи в экземпляр SAT.

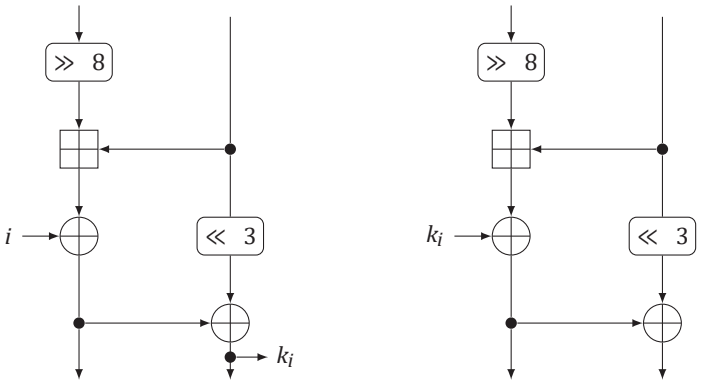
В общем случае и SAT, и SMT трудноразрешимы. Как и в случае смешанно-целочисленного линейного программирования, мы примем как данность существование решателей, способных практически разрешать такие задачи.

3.3.1. Пример: шифр add-rotate-xor

Некоторые шифры для введения нелинейности полагаются не на маленькие S-блоки, а на операции с длинными операндами, которые аппаратно поддерживаются современными процессорами (например, сложение по модулю или поразрядное И). В этом разделе приведен пример шифра типа add-rotate-xor (ARX, «сложение – циклический сдвиг – исключающее или»). Оптимизацию следов в таких шифрах часто удобно моделировать как SMT-задачу с битовыми векторами в качестве переменных.

Раундовая функция для этого примера показана на рис. 3.4b. Символ $\boxplus$ представляет целочисленное сложение по модулю 2^n, где $n \in \{24, 32, 48, 64\}$. Раундовые ключи генерируются аналогичной функцией, показанной на рис. 3.4a. Этот шифр называется Speck и был спроектирован в Агентстве национальной безопасности США.

Как и для шифра типа Rijndael из раздела 3.2.1, количество раундов пока задавать не будем. Но в отличие от шифра типа Rijndael для Speck не существует простого рассуждения, которое позволило бы ограничить абсолютную корреляцию следов сверху.



(a) Один раунд развертки ключа (b) Один раунд шифра

Рис. 3.4. Шифр Speck типа ARX

3.3.2. Построение модели

Для элементов корреляционных матриц линейных функций шифра (циклических сдвигов и ИСКЛЮЧАЮЩЕГО ИЛИ двух ветвей) имеются замкнутые формулы, которые можно без труда преобразовать в ограничения на битовые векторы. Основную трудность представляет сложение по модулю степени двойки. Однако оказывается, что для элементов корреляционной матрицы этой операции также существует простая формула.

Эта формула выводится путем преобразования графика функции сложения по модулю в график функции, с которой проще работать. Графиком функции $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ является множество пар $G_\mathsf{F} = \{(x, \mathrm{F}(x)) \mid x \in \mathbb{F}_2^n\}$. В частности, $G_{\boxplus}$ является графиком функции сложения по модулю 2^n $x \parallel y \mapsto x \boxplus y$.

Лемма 3.2 (Шульте–Гееса). *Пусть* $M \in \mathbb{F}_2^{n\times n}$ *– нижнетреугольная матрица*

$$M = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{bmatrix}.$$

Пусть также $\mathsf{Q} : \mathbb{F}_2^{2n} \to \mathbb{F}_2^n$ *– функция, определенная как* $\mathsf{Q}(x \parallel y) = M(x \wedge y)$, *где* $x \wedge y$ *– поразрядное И величин* x *и* y. *Отображение* $(x \parallel y, z) \mapsto ((x + z) \parallel (y + z), x + y + z)$ *является биекцией* $G_{\boxplus}$ *в* G_Q.

Доказательство. Нетрудно видеть, что отображение $(x \parallel y, z) \mapsto ((x + z) \parallel (y + z), x + y + z)$ – биекция, если оно корректно определено. Поэтому достаточно проверить, что если $z = x \boxplus y$, то

$$x + y + z = \mathsf{Q}(x + z \parallel y + z).$$

Однако $c = x + y + z$ равно вектору битов переноса, полученных при сложении x и y по модулю. Эти биты переноса c удовлетворяют соотношениям $c_1 = 0$ и

$$c_{i+1} = c_i + (x_i + z_i)(y_i + z_i).$$

Эти соотношения следуют из школьного алгоритма сложения. Отсюда вектор переноса c равен $M((x + z)\wedge(y + z) = \mathsf{Q}(x + z \parallel y + z)$. □

В теореме 3.3 знак $\preccurlyeq$ обозначает поэлементное упорядочение битовых векторов длины n, определенное как $0 \preccurlyeq 0$, $0 \preccurlyeq 1$ и $1 \preccurlyeq 1$. Заметим, что $x \preccurlyeq y$ – то же самое, что «x_i влечет y_i» для $i = 1, \ldots, n$. Следовательно, $x \preccurlyeq y$ можно также переписать в виде $x \wedge \bar{y} = 0$, где $\bar{y}$ – поразрядное дополнение y.

Теорема 3.3. *Пусть* u, v *и* w *– маски в* $\mathbb{F}_2^n$. Тогда $C^{\boxplus}_{w,u\parallel v} \neq 0$ *тогда и только тогда, когда* $(u + w) \vee (v + w) \preccurlyeq M^\mathsf{T}(u + v + w)$. *Кроме того, в этом случае*

$$C^{\boxplus}_{w,\,u\parallel v} = (-1)^{(u+w)^\mathsf{T}(v+w)}/2^{\mathsf{wt}(M^\mathsf{T}(u+v+w))}.$$

Доказательство. Корреляционная матрица функции $\boxplus\colon \mathbb{F}_2^{2n} \to \mathbb{F}_2^n$ удовлетворяет соотношению

$$C^{\boxplus}_{w,u\|v} = \frac{1}{2^n} \sum_{(x\|y,z)\in G_{\boxplus}} (-1)^{u^\intercal x + v^\intercal y + w^\intercal z}.$$

Применяя лемму 3.2 и подстановку, получаем

$$\begin{aligned} C^{\boxplus}_{w,u\|v} &= \frac{1}{2^n} \sum_{(x\|y,z)\in G_{\mathsf{Q}}} (-1)^{u^\intercal (x+z) + v^\intercal (y+z) + w^\intercal (x+y+z)} \\ &= \frac{1}{2^n} \sum_{(x\|y,z)\in G_{\mathsf{Q}}} (-1)^{(u+w)^\intercal x + (v+w)^\intercal y + (u+v+w)^\intercal z} \\ &= C^{\mathsf{Q}}_{u+v+w,(u+w)\|(v+w)}. \end{aligned}$$

Поскольку Q с точностью до умножения на M совпадает с поразрядным И, результат следует из формулы для корреляционной матрицы поразрядного И, приведенной в упражнении 2.8. □

В качестве примера построим модель одного раунда шифра Speck. Если входная маска равна $u_1 \| u_2$, а выходная – $v_1 \| v_2$, то входная маска сложения по модулю равна $(u_1 \gg 8) \| (u_2 + (v_2 \gg 3))$. Выходная маска равна $v_1 + v_2$. Подстановка в теорему 3.3 дает ограничения на битовый вектор, при которых одноразундовая линейная аппроксимация эффективна. Комбинирование этих условий дает ограничения, при которых след имеет ненулевую корреляцию. Наконец, следует добавить еще одно ограничение, потребовав, чтобы сумма весов линейной аппроксимации (которую также дает теорема 3.3) была равна константе W. Если существует след с абсолютной корреляцией 2^{-W}, то решатель его найдет. В противном случае он сообщит, что задача невыполнима.

Большинство операций, имеющих место при вышеупомянутых ограничениях, предоставлены SMT-решателями, которые поддерживают битовые векторы в качестве переменных. Основные исключения – умножение на M и вес Хэмминга. Однако их легко выразить в терминах операций над битовыми векторами и целочисленного сложения.

3.3.3. Решение модели

После того как модель построена, ее можно решить, положив вес следа W равным нижней границе (в худшем случае 0). Если решатель сообщает, что задача невыполнима, то производится попытка решить ее снова с весом следа $W + 1$. Этот процесс повторяется, пока не будет найдено решение с минимальным весом.

К числу популярных SMT-решателей с поддержкой битовых векторов относятся Boolector и Z3. Задачи формулируются либо с помощью зависящего от решателя API, либо в широко поддерживаемом файловом формате LibSMT.

3.4. Историческая справка

Методы автоматизации поиска линейных следов применялись с первых дней линейного криптоанализа, начиная с метода ветвей и границ Мацуи. Примерно в 2010 году популярной альтернативой стали готовые решатели MILP и SAT, а к настоящему времени накопилась обширная литература по этому вопросу.

Стратегия широкого следа была предложена Даменом в его докторской диссертации, а впоследствии развита в совместной работе с Рэйменом. Блочный шифр Rijndael, лежащий в основе примера из раздела 3.2.1, был спроектирован Даменом и Рэйменом в 1997 году, а его 128-битовая версия была стандартизована Национальным институтом стандартов и технологий США (NIST) в 2001 году.

Хотя терминология «шифр типа ARX» появилась позже, первые шифры, спроектированные на основе этого подхода, были опубликованы еще в 1980-х годах. Блочный шифр Speck, использованный в качестве примера в разделе 3.3.1, появился в 2013 году. Эффективный алгоритм вычисления корреляций линейных аппроксимаций для сложения по модулю впервые описал Уоллрен в 2003 году. Упрощенная формула в теореме 3.3 принадлежит Шульте–Геесу.

3.5. Литература

Beaulieu, Ray et al. (2013). The SIMON and SPECK Families of Lightweight Block Ciphers. Cryptology ePrint Archive, Report 2013/404. url: *https://eprint.iacr.org/2013/404*.

Daemen, Joan and Vincent Rijmen (Dec. 2001). «The Wide Trail Design Strategy». In: 8th IMA International Conference on Cryptography and Coding. Ed. by Bahram Honary. Vol. 2260. LNCS. Springer, Berlin, Heidelberg, pp. 222–238. doi: 10.1007/3-540-45325-3_20.

Schulte-Geers, Ernst (2013). «On CCZ-equivalence of addition mod 2^n». In: Designs, Codes and Cryptography 66, pp. 111–127.

Wallen, Johan (Feb. 2003). «Linear Approximations of Addition Modulo 2^n». In: FSE 2003. Ed. by Thomas Johansson. Vol. 2887. LNCS. Springer, Berlin, Heidelberg, pp. 261–273. doi: 10.1007/978-3-540-39887-5_20.

3.6. Упражнения

Упражнение 3.1

В этом упражнении исследуются свойства шага `MixColumns` демонстрационного шифра типа Rijndael. В разделе 3.2.1 было отмечено, что это отображение соответствует блочной матрице, состоящей из блоков 3×3 над $\mathbb{F}_2$.

$$M = \begin{bmatrix} 0 & I & I & I \\ I & 0 & I & I \\ I & I & 0 & I \\ I & I & I & 0 \end{bmatrix}.$$

1. Покажите, что M обратима, и найдите обратную ей матрицу.
2. Чему равно число линейных разветвлений M?

* Упражнение 3.2

Рассмотрим демонстрационный шифр типа Rijndael из раздела 3.2.1. Покажите, что

1) абсолютная корреляция любого двухраундового линейного следа не превышает $1/2^4$;
2) абсолютная корреляция любого четырехраундового линейного следа не превышает $1/2^{16}$.

Упражнение 3.3

В этом упражнении исследуются некоторые свойства числа разветвлений $\mathcal{B}(M)$ матрицы M или линейного отображения $x \mapsto Mx$ над $\mathbb{F}_2$ (см. определение 3.1). Предположим, что M – матрица размера $bm \times bn$ и что вес Хэмминга определен относительно b-битовых блоков.

1. Докажите, что если M обратима, то $\mathcal{B}(M) = \mathcal{B}(M^{-1})$.
2. Используя определение числа разветвлений, покажите, что вес Хэмминга каждого элемента векторного пространства $C = \{u \,\|\, v \mid v \in \mathbb{F}_2^{bm}$ и $u = M^{\mathrm{T}}v\}$ не меньше $\mathcal{B}(M)$. Какое векторное пространство C соответствует M^{-1}?
3. Покажите, что число разветвлений M не превышает $n + 1$.

В литературе по теории кодирования векторное пространство C называется *блочным кодом*. Верхняя граница $n + 1$ числа разветвлений называется *границей Синглтона*. Матрица M с числом разветвлений $n + 1$ называется *матрицей с максимальным разделением*, или MDS-матрицей (maximum distance separable).

* Упражнение 3.4

В этом упражнении вводится конкретное построение MDS-матриц, основанное на полиномах над конечными полями. Поэтому для его решения необходимо знакомство с теорией конечных полей. Как было сказано в упражнении 3.3, матрица размера $bn \times bn$ над полем $\mathbb{F}_2$, состоящая из $b \times b$ блоков, называется MDS-матрицей, если ее число разветвлений $\mathcal{B}(M)$ равно $n + 1$.

1. Пусть $\mathbb{F}_{2^b}$ – конечное поле порядка 2^b и L: $\mathbb{F}_{2b}^n \to \mathbb{F}_{2b}^n$ – произвольное линейное отображение. Покажите, что существует обратимое $\mathbb{F}_2$-линейное отображение $\beta : \mathbb{F}_{2b}^n \to \mathbb{F}_{2b}^n$ и матрица M размера $bn \times bn$ над $\mathbb{F}_2$ такие, что $x \mapsto \beta^{-1}(M^{\mathrm{T}}\beta(x))$ равно L. Покажите, что число разветвлений M удовлетворяет соотношению

$$\mathcal{B}(M) = \max_{x \neq 0} \mathrm{wt}(x) + \mathrm{wt}(L(x)),$$

   где $\mathrm{wt}(x)$ – число ненулевых элементов $x \in \mathbb{F}_{2b}^n$.

2. Пусть p_x – полином степени $\leqslant n - 1$ с коэффициентами $x_1, \ldots, x_n$, где $x_1, \ldots, x_n$ – элементы вектора $x \in \mathbb{F}_{2b}^n$. Покажите, что для любых $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_{2b}^n$ следующее отображение L: $\mathbb{F}_{2b}^n \to \mathbb{F}_{2b}^n$ является линейным:

$$L : x \mapsto \begin{bmatrix} p_x(\alpha_1) \\ \vdots \\ p_x(\alpha_n) \end{bmatrix}.$$

3. Постройте $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_{2^b}$ так, чтобы матрица M, построенная по функции L, имела число разветвлений $n + 1$.

Код, соответствующий матрице M (см. упражнение 3.3), называется кодом Рида–Соломона.

Упражнение 3.5

S-блок в блочном шифре Rijndael основан на отображении $x \mapsto 1/x$ в конечном поле $\mathbb{F}_{2^n}$. Пусть $\beta : \mathbb{F}_{2^n} \to \mathbb{F}_2^n$ – обратимое линейное отображение, A – матрица размера $n \times n$ над $\mathbb{F}_2$, а b – вектор, принадлежащий $\mathbb{F}_2^n$. Определим S-блок $\mathsf{S}: \mathbb{F}_2^n \to \mathbb{F}_2^n$ как

$$\mathsf{S}(x) = \begin{cases} A\beta^{-1}(1/\beta(x)) + b, & \text{если } x \neq 0, \\ b & \text{в противном случае.} \end{cases}$$

Цель этого упражнения – показать, что все нетривиальные линейные аппроксимации шифра AES имеют низкую абсолютную корреляцию. Для доказательства этого результата можно воспользоваться следующей оценкой *сумм Клоостермана*:

$$\left| \sum_{x \in \mathbb{F}_{2^n}^\times} (-1)^{\mathrm{Tr}(x + c/x)} \right| \leq 2^{n/2+1},$$

где $\mathrm{Tr} : \mathbb{F}_{2^8} \to \mathbb{F}_2$ – функция следа, а $c \in \mathbb{F}_{2^n}$ – константа. Для решения этого упражнения необходимо знакомство с теорией конечных полей.

1. Покажите, что достаточно найти верхнюю границу корреляций линейных аппроксимаций функции, определенной как $x \mapsto \beta^{-1}(1/\beta(x))$ для $x \neq 0$ и $0 \mapsto 0$.
2. Докажите, что абсолютная корреляция всех нетривиальных линейных аппроксимаций S не превышает $2^{1-n/2}$.
3. Найдите описание S-блока шифра Rijndael и вычислите его корреляционную матрицу. Сравните со своим результатом для $n = 8$.

Упражнение 3.6

Используйте метод ветвей и границ, смешанно-целочисленное линейное программирование (MILP) или SMT, чтобы автоматизировать оптимизацию следов в демонстрационном шифре из раздела 1.1. Для решения этого упражнения воспользуйтесь одним из следующих инструментов:

- любым языком программирования для реализации метода ветвей и границ;
- Python с пакетом Google OR-Tools[1] для создания и решения MILP-моделей;
- Python с пакетом PySMT[2] для создания и решения SMT-моделей.

[1] *https://developers.google.com/optimization*.

[2] *https://github.com/pysmt/pysmt*.

Используя свою модель демонстрационного шифра, решите следующие задачи.

1. Проверить результаты из примера 2.3.
2. Найти линейный след по пяти раундам с корреляцией $\pm 2^{-5}$.
3. Выбрать один из линейных следов из предыдущей задачи и найти все линейные следы для той же аппроксимации.

Упражнение 3.7

Воспользуйтесь MILP, чтобы смоделировать распространение линейных следов в демонстрационном шифре типа Rijndael из раздела 3.2.1.

1. Из упражнения 3.2 следует, что максимальная абсолютная корреляция линейного следа по четырем раундам не превышает 2^{-16}. Найдите след, для которого эта оценка реализуется.
2. Для найденной выше линейной аппроксимации найдите линейный след (или следы) с предыдущей по величине абсолютной корреляцией. Обсудите, какие следствия вытекают из этих дополнительных следов.

Упражнение 3.8

Воспользуйтесь SMT, чтобы смоделировать распространение линейных следов в демонстрационном шифре Simon с размером блока 32 бита, описанном в упражнении 2.9.

1. Проверьте результат, найденный вами в упражнении 2.9.
2. Существуют ли еще какие-то линейные следы для той же линейной аппроксимации, которые следует принимать во внимание?

* Упражнение 3.9

Воспользуйтесь SMT, чтобы смоделировать распространение линейных следов в шифре Speck с размером блока 64 бита, описанном в разделе 3.3.1. Найдите оптимальный линейный след по семи раундам. Реализуйте первый алгоритм Мацуи, чтобы восстановить один бит информации о секретном ключе, и проверьте, работает ли он в соответствии с ожиданиями.

Глава 4

Статистика линейного криптоанализа

Определение эффективности линейного криптоанализа – это приложение статистической теории. В этой главе мы рассмотрим некоторые базовые понятия статистики и обсудим, как они используются для оценивания стоимости линейных атак и второго алгоритма Мацуи в частности.

4.1. Статистический вывод

Параметрическое оценивание и проверка гипотез – две тесно связанные задачи статистического вывода. Обе они важны для анализа линейных атак, поэтому в настоящем разделе мы дадим обзор базовых принципов. Основным результатом является теорема 4.1 о проверке гипотезы о совпадении двух нормальных распределений с одинаковой дисперсией. Она неоднократно используется в этой книге. В приложении А приведены необходимые сведения о нормальном распределении.

4.1.1. Статистические оценки

Пусть $\mathbf{x}$ – случайная величина с распределением вероятностей P_θ, где θ – неизвестный параметр. Например, предположим, что $\mathbf{x}$ имеет нормальное распределение со средним μ и дисперсией σ^2, но мы не знаем значения μ. Для краткости обозначим это условие $\mathbf{x} \sim \mathcal{N}(\mu, \sigma^2)$.

Неизвестный параметр θ можно оценить, или «вывести», на основе *выборки* $x_1, \ldots, x_q$ из распределения P_θ. *Оценкой* (статистического) параметра θ распределения P_θ называется функция f, которая отображает выборку $(x_1, \ldots, x_q)$ на величину оценки параметра θ.

Пример 4.1. Типичной оценкой среднего $\mu = \mathbb{E}(\mathbf{x})$ случайной величины $\mathbf{x}$ является *выборочное среднее*

$$\widehat{\mu}(x_1, \ldots, x_q) = \frac{1}{q} \sum_{i=1}^{q} x_i.$$

Если среднее распределения $\mathbf{x}$ – неизвестный параметр, то эту оценку можно использовать для вывода его значения. ▷

Для обсуждения статистических свойств оценок мы будем рассматривать саму выборку как случайную величину. Распределение случайной выборки $(\mathbf{x}_1, \ldots, \mathbf{x}_q)$ зависит от распределения P_θ, а также от стратегии формирования выборки. Простейшая стратегия – выборка с возвращением: в этом случае $\mathbf{x}_1, \ldots, \mathbf{x}_q$ независимы и все имеют распределение P_θ. Оценка f параметра θ называется несмещенной, если

$$\mathop{\mathbb{E}}_{\mathbf{x}_1,\ldots,\mathbf{x}_q} f(\mathbf{x}_1, \ldots, \mathbf{x}_q) = \theta.$$

Быть может, вопреки интуиции смещенные оценки иногда бывают полезны, в частности когда $f(\mathbf{x}_1, \ldots, \mathbf{x}_q)$ с большой вероятностью близко к θ.

Пример 4.2. Выборочное среднее, определенное в примере 4.1, является несмещенной оценкой среднего случайной величины $\mathbf{x}$. В самом деле, пусть $(\mathbf{x}_1, \ldots, \mathbf{x}_q)$ – случайная выборка, так что частное (маргинальное) распределение $\mathbf{x}_i$ совпадает с распределением $\mathbf{x}$ для $i = 1, \ldots, q$. Поскольку

$$\mathop{\mathbb{E}}_{\mathbf{x}_1,\ldots,\mathbf{x}_q} \widehat{\mu}(\mathbf{x}_1, \ldots, \mathbf{x}_q) = \mathop{\mathbb{E}}_{\mathbf{x}_1,\ldots,\mathbf{x}_q} \frac{1}{q} \sum_{i=1}^{q} \mathbf{x}_i = \frac{1}{q} \sum_{i=1}^{q} \mathbb{E}(\mathbf{x}_i) = \mathbb{E}(\mathbf{x}),$$

выборочное среднее является несмещенной оценкой среднего $\mathbf{x}$. ▷

Помимо среднего оценки, мы должны учитывать, насколько сильно оценка отклоняется от своего среднего. Одной из мер такого отклонения является дисперсия

$$\mathop{\mathbb{V}}_{\mathbf{x}_1,\ldots,\mathbf{x}_q} f(\mathbf{x}_1, \ldots, \mathbf{x}_q) = \mathop{\mathbb{E}}_{\mathbf{x}_1,\ldots,\mathbf{x}_q} \Big(f(\mathbf{x}_1, \ldots, \mathbf{x}_q) - \mu \Big)^2,$$

где μ – среднее оценки. Если оценка f несмещенная, то $\mu = \theta$.

Пример 4.3. Обозначим σ^2 дисперсию $\mathbf{x}$. Если образцы $\mathbf{x}_1, \ldots, \mathbf{x}_q$ являются независимыми случайными величинами с таким же частным распределением, как у $\mathbf{x}$, то дисперсия выборочного среднего равна

$$\mathop{\mathbb{V}}_{\mathbf{x}_1,\ldots,\mathbf{x}_q} \widehat{\mu}(\mathbf{x}_1, \ldots, \mathbf{x}_q) = \mathop{\mathbb{V}}_{\mathbf{x}_1,\ldots,\mathbf{x}_q} \frac{1}{q} \sum_{i=1}^{q} \mathbf{x}_i = \frac{1}{q^2} \sum_{i=1}^{q} \mathbb{V}(\mathbf{x}_i) = \frac{\sigma^2}{q}.$$

Третье равенство следует из того, что дисперсия суммы независимых случайных величин равна сумме дисперсий слагаемых.

4.1.2. Проверка гипотез

Цель *проверки статистической гипотезы* – доказать ложность гипотезы о распределении вероятностей при наличии неопределенностей. По соглашению, гипотеза, ложность которой следует установить, называется нулевой гипотезой. Ее часто сравнивают со второй гипотезой, которая называется альтернативной. Это классическая интерпретация проверки статистических гипотез, введенная в обиход Фишером. В криптоанализе более естественно следовать интерпретации Неймана–Пирсона, при которой проверка статистических

гипотез рассматривается как задача принятия решения (или «различения»). В таком случае альтернативная гипотеза считается конкурирующей.

Проверка гипотезы принимает на входе значение статистики критерия, а на выходе сообщает, следует ли отвергнуть нулевую гипотезу. Статистика критерия – это функция наблюдаемых данных (выборки). Например, для проверки гипотезы о параметре θ распределения P_θ статистикой критерия могла бы быть оценка θ.

В этой главе предполагается, что статистика критерия f является вещественной функцией и что проверка гипотезы заключается в сравнении $t = f(x_1, \ldots, x_q)$ с пороговым значением τ. Если $t \geqslant \tau$, то проверка не опровергает нулевую гипотезу («принимает» ее). Если же $t < \tau$, то нулевая гипотеза отвергается. Обозначим $\mathbf{t}_{\text{null}}$ значение статистики критерия для случайной выборки, полученное в предположении нулевой гипотезы. Аналогично обозначим $\mathbf{t}_{\text{alt}}$ значение статистики критерия для случайной выборки, полученное в предположении альтернативной гипотезы. Для любой проверки гипотезы важны две вероятности:

$$P_S = \Pr[\mathbf{t}_{\text{null}} \geqslant \tau],$$
$$P_F = \Pr[\mathbf{t}_{\text{alt}} \geqslant \tau].$$

Вероятность P_S называется *вероятностью истинно положительного результата*, или (в криптоанализе) *вероятностью успеха*, а P_F – *вероятностью ложноположительного результата*. В общем случае всегда существует взаимосвязь между P_S и $1 - P_F$, потому что при уменьшении τ увеличивается как P_S, так и P_F (но обычно на разные величины). Это означает, что ни P_S, ни $1 - P_F$ не могут служить хорошими мерами качества, если рассматривать их по отдельности. Криптографы иногда используют следующую меру качества, которая называется *преимуществом*.

$$|P_S - P_F|.$$

Заметим, что это определение симметрично относительно нулевой и альтернативной гипотез. Понятие преимущества выбрано произвольно. Другие меры обсуждаются в главе 7.

Далее в этом разделе мы изучим проверку статистических гипотез на примере среднего нормального распределения. В этом случае нулевая гипотеза утверждает, что распределение статистики критерия является нормальным со средним $\mu \neq 0$, а альтернативная – что это нормальное распределение с нулевым средним. Обе гипотезы утверждают, что стандартное отклонение равно $\sigma/\sqrt{q}$. Как обсуждалось в примере 4.3, это стандартное отклонение выборочного среднего q образцов от распределения со стандартным отклонением σ. Ситуация в целом показана на рис. 4.1.

Следующий результат важен для анализа линейных атак. Он используется на протяжении всей книги. В теореме 4.1 Ф обозначает функцию распределения нормального распределения с нулевым средним и единичной дисперсией.

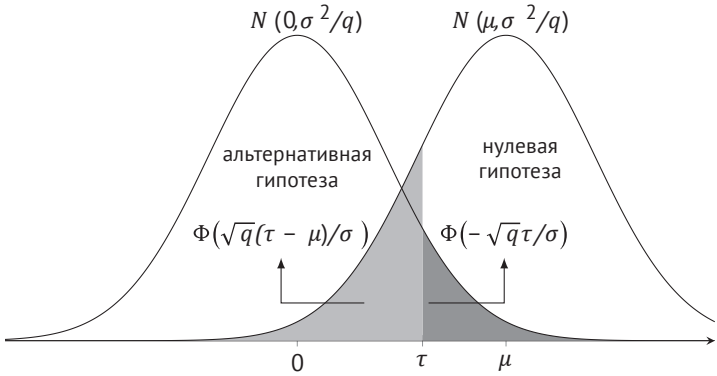


Рис. 4.1. Распределение при нулевой и альтернативной гипотезах

Теорема 4.1. *Для проверки гипотезы, в которой выборка сравнивается с порогом, чтобы различить два нормальных распределения с одинаковой дисперсией σ^2/q и средним $\mu \neq 0$ (нулевая гипотеза) и $\mu = 0$ (альтернативная гипотеза), вероятность успеха P_S и вероятность ложноположительного результата $P_F \leqslant P_S$ удовлетворяют соотношению*

$$q = \left(\frac{\Phi^{-1}(P_S) - \Phi^{-1}(P_F)}{\mu/\sigma} \right)^2.$$

Доказательство. Предположим, что $\mu > 0$. Похожее рассуждение проходит и для $\mu < 0$. Как показано на рис. 4.1, проверка принимает нулевую гипотезу, когда статистика критерия (равная самому выборочному значению) превосходит τ. Как и выше, обозначим $\mathbf{t}_{null}$ значение статистики критерия для случайной выборки в предположении истинности нулевой гипотезы, а $\mathbf{t}_{alt}$ – ее значение для случайной выборки в предположении альтернативной гипотезы.

В предположении истинности нулевой гипотезы статистика критерия имеет нормальное распределение со средним μ и дисперсией σ^2/q. Вероятность успеха равна вероятности того, что $\mathbf{t}_{null}$ больше τ, когда нулевая гипотеза истинна:

$$P_S = \Pr\left[\mathbf{t}_{null} \geq \tau\right] = 1 - \Phi\left((\tau - \mu)/(\sigma/\sqrt{q})\right) = \Phi\left(\sqrt{q}\,(\mu - \tau)/\sigma\right).$$

Отсюда $\sqrt{q}\tau/\sigma = \sqrt{q}\mu/\sigma - \Phi^{-1}(P_S)$. Вероятность ложноположительного результата – это вероятность того, что $\mathbf{t}_{alt}$ больше τ, когда истинна альтернативная гипотеза:

$$P_F = \Pr\left[\mathbf{t}_{alt} \geq \tau\right] = 1 - \Phi\left(\tau/(\sigma/\sqrt{q})\right) = \Phi\left(-\sqrt{q}\,\tau/\sigma\right).$$

Эквивалентно $\sqrt{q}\tau/\sigma = -\Phi^{-1}(P_F)$. Из этих двух выражений для $\sqrt{q}\tau/\sigma$ получаем

$$\Phi^{-1}(P_S) = \sqrt{q}\,\mu/\sigma + \Phi^{-1}(P_F).$$

Изменение порядка членов и возведение в квадрат дает искомый результат при условии, что $\Phi^{-1}(P_S) \geqslant \Phi^{-1}(P_F)$. Поскольку Φ строго возрастает, последнее условие эквивалентно $P_S \geqslant P_F$. □

4.2. Восстановление ключа с помощью проверки статистических гипотез

Напомним (см. раздел 1.4.2), что алгоритм 2 Мацуи вычисляет эмпирическую корреляцию линейной аппроксимации внутренней части шифра для всех возможных значений релевантных битов ключа в его внешней части. Этот подход подразумевает, что эмпирическая корреляция больше (по абсолютной величине) для правильного ключа, чем для неправильных.

С точки зрения проверки статистических гипотез, для каждой эмпирической корреляции, вычисленной алгоритмом 2 Мацуи, мы должны решить, является ли соответствующий ключ правильным. Если существует K возможных ключей, то, чтобы оставить наиболее многообещающих кандидатов, потребуется K проверок гипотезы. Предположим, что для всех этих проверок вероятность успеха равна P_S, а вероятность ложноположительного результата равна P_F. По определению, вероятность восстановления правильного ключа равна P_S. Кроме того, среднее число ключей-кандидатов равно $P_S + P_F(K - 1) \approx P_F K$.

В следующих двух разделах анализируется число известных открытых текстов, необходимое алгоритму 2 Мацуи, чтобы с вероятностью P_S восстановить правильный частичный ключ как один из приблизительных $P_F K$ ключей-кандидатов. С точностью до обсуждаемых ниже аппроксимаций это то же самое, что информационная сложность различения с вероятностью успеха P_S и вероятностью ложноположительного результата P_F на основе линейной аппроксимации внутренней части шифра. В разделе 4.2.1 рассматривается случай, когда линейная аппроксимация имеет известную (не зависящую от ключа) корреляцию. Случай неизвестной корреляции обсуждается в разделе 4.2.2.

4.2.1. Известная корреляция

Для криптоанализа типична ситуация, когда аналитику точно неизвестны распределения, которые проверка гипотез должна различить. Чтобы можно было проанализировать стоимость линейных различителей, мы предложим *модель*, т. е. предположим конкретные распределения и выполним некоторые дополнительные аппроксимации. В этой главе мы начнем с довольно грубой модели, которая называется «простой». Уточнение простой модели обсуждается в главе 7.

В простой модели нулевая гипотеза утверждает, что корреляция линейной аппроксимации $c \neq 0$. Альтернативная гипотеза утверждает, что корреляция *в точности равна нулю*. Корни последней гипотезы лежат в предположении, что неправильные ключи должны давать эмпирическую корреляцию, более близкую к нулю; это предположение известно под названием «гипотеза рандомизации с неправильным ключом». Интуитивно понятно, что неправильная догадка «рандомизирует» статистику критерия, что должно приводить к корреляции, близкой к нулю.

В главе 7 будет показано, что даже линейные аппроксимации равномерно распределенной случайной перестановки или функции редко имеют корреляцию, в точности равную нулю. Следовательно, простая модель необязательно является точным отображением реальности, а полученные с ее помощью результаты следует трактовать как аппроксимации.

В простой модели делаются следующие дополнительные технические предположения.

Малая корреляция: квадрат корреляции пренебрежимо мал по сравнению с единицей.

Большие данные: количество данных q достаточно велико, так что, например, приближение биномиального распределения нормальным является точным.

Модель выборки: выборка открытых текстов производится случайным и равномерным образом *с возвращением*. Отсюда, в частности, следует, что образцы независимы.

Обычно эти предположения реалистичны, но при необходимости от них можно отказаться без особых трудностей (см., например, упражнение 4.1).

В простой модели информационная сложность линейного различителя описывается, по существу, теоремой 4.1. Для равномерного случайного входа $\mathbf{x}$ и соответствующего выхода $\mathbf{y} = \mathsf{F}(\mathbf{x})$ определим случайную величину

$$\mathbf{z} = (-1)^{u^\mathsf{T}\mathbf{x}+v^\mathsf{T}\mathbf{y}}.$$

По определению, $\mathbb{E}(\mathbf{z})$ – корреляция линейной аппроксимации (u,v) функции F. Если истинна нулевая гипотеза, то корреляция равна c. Если же истинна альтернативная гипотеза, то она равна нулю. Кроме того, $\mathbb{E}(\mathbf{z}^2) = 1$ для любой гипотезы. Следовательно, для нулевой гипотезы $\mathbb{V}(\mathbf{z}) = 1 - c^2 \approx 1$, а для альтернативной $\mathbb{V}(\mathbf{z}) = 1$.

Набор q известных открытых текстов и соответствующих им шифртекстов, иначе говоря – значений $\mathbf{z}_1 = (-1)^{u^\mathsf{T}\mathbf{x}_1+v^\mathsf{T}\mathbf{y}_1}$, ..., $\mathbf{z}_q = (-1)^{u^\mathsf{T}\mathbf{x}_q+v^\mathsf{T}\mathbf{y}_q}$, соответствует выборке $\mathbf{z}$.. Выполняемые нами статистические проверки основаны на выборочном среднем или *эмпирической корреляции*

$$\widehat{\mathbf{c}} = \frac{1}{q}\sum_{i=1}^{q}\mathbf{z}_i.$$

Как обсуждалось в разделе 4.1.1, выборочное среднее является оценкой среднего. Иными словами, эмпирическая корреляция – это статистическая оценка корреляции.

Поскольку выборка из входов производится с возвращением, случайные образцы $\mathbf{z}_1$, ..., $\mathbf{z}_q$ независимы при обеих гипотезах. Если q велико, то из центральной предельной теоремы (теорема А.1 в приложении А) следует, что $\mathbf{c}$ приближенно совпадает с нормальным распределением. В силу примеров 4.2 и 4.3, среднее $\hat{\mathbf{c}}$ равно истинному среднему по генеральной совокупности, т. е. 0 или c, а дисперсия $\hat{\mathbf{c}}$ равна дисперсии по генеральной совокупности, поделенной на размер выборки q, т. е. $1/q$. Поэтому из теоремы 4.1 следует такой результат.

Следствие 4.2. *Информационная сложность линейного различителя в простой модели, где используется линейная аппроксимация с известной корреляцией c, равна*

$$q = \left(\frac{\Phi^{-1}(P_\mathsf{S}) - \Phi^{-1}(P_\mathsf{F})}{c}\right)^2,$$

где P_S *– вероятность успеха, а* $P_F \leqslant P_S$ *– вероятность ложноположительного результата.*

Доказательство. Достаточно положить $\mu = c$ и $\sigma = 1$ в теореме 4.1. □

4.2.2. Неизвестная корреляция

Для большинства атак, описанных в литературе, корреляция линейной аппроксимации зависит от (неизвестного) ключа. Кроме того, часто линейные аппроксимации бывает трудно вычислить точно, даже когда ключ известен, потому что в корреляцию могут вносить вклад много линейных следов.

Когда корреляция неизвестна, основная трудность применения критерия из раздела 4.2.1 заключается в том, что неизвестен знак корреляции. Одна из возможных стратегий – отдельно рассмотреть случаи положительной и отрицательной корреляций, выполнив две проверки гипотез для каждого возможного ключа. Это увеличивает вероятность ложноположительного результата, поскольку количество ключей-кандидатов приблизительно удваивается. Другая стратегия – работать со статистикой критерия, не зависящей от знака c. Например, можно разработать критерий, основанный на $|\hat{\mathbf{c}}|$ или, что эквивалентно, на $\hat{\mathbf{c}}^2$. Нулевая гипотеза принимается, если абсолютная величина или квадрат корреляции превосходит порог τ. Этот критерий работает при условии, что абсолютная корреляция при нулевой гипотезе значительно больше, чем при альтернативной.

Следующая теорема дает вероятность успеха, когда $|c|$ известна, но знак c неизвестен. Для шифров с чередованием ключа это соответствует случаю линейных аппроксимаций, в которых доминирует корреляция одного линейного следа. Общий случай, когда $|c|$ также зависит от ключа, обсуждается позже. В теореме 4.3 делаются такие же предположения, как в простой модели, с тем отличием, что проверка основана на абсолютном значении или квадрате эмпирической корреляции.

Теорема 4.3. *Вероятность успеха линейного различителя в простой модели, основанной на абсолютной величине или квадрате эмпирической корреляции и использующей линейную аппроксимацию с известной абсолютной корреляцией* $|c|$*, равна*

$$P_S = \Phi\left(\Phi^{-1}(P_F/2) + |c|\sqrt{q}\right) + \Phi\left(\Phi^{-1}(P_F/2) - |c|\sqrt{q}\right),$$

где q *– информационная сложность, а* P_F *– вероятность ложноположительного результата.*

Доказательство. Эмпирическая корреляция $\hat{\mathbf{c}}_{\text{null}}$ при нулевой гипотезе имеет нормальное распределение со средним c и дисперсией $1/q$. Аналогично эмпирическая корреляция $\hat{\mathbf{c}}_{\text{alt}}$ при альтернативной гипотезе имеет нулевое среднее и дисперсию $1/q$.

Поскольку $\hat{\mathbf{c}}^2 \geqslant \tau$ эквивалентно $|\hat{\mathbf{c}}| \geqslant \sqrt{\tau}$ для $\tau \geqslant 0$, не имеет значения, будем ли мы использовать абсолютную величину корреляции или ее квадрат. Следовательно, для $\tau \geqslant 0$ вероятность ложноположительного результата P_F равна

$$P_{\mathsf{F}} = \Pr\left[|\widehat{\mathfrak{c}}_{\mathsf{alt}}| \geq \tau\right] = 2\Phi\big(-\sqrt{q}\tau\big).$$

Решение этого уравнения относительно τ дает $\sqrt{q}\tau = -\Phi^{-1}(P_{\mathsf{F}}/2)$. Вероятность успеха P_{S} равна

$$\begin{aligned} P_{\mathsf{S}} &= \Pr\left[|\widehat{\mathfrak{c}}_{\mathsf{null}}| \geq \tau\right] \\ &= \Pr\left[\widehat{\mathfrak{c}}_{\mathsf{null}} \geq \tau\right] + \Pr\left[\widehat{\mathfrak{c}}_{\mathsf{null}} \leq -\tau\right] \\ &= \Phi\big(-\sqrt{q}(\tau - c)\big) + \Phi\big(-\sqrt{q}(\tau + c)\big). \end{aligned}$$

Поскольку это выражение является четной функцией от c, его можно переписать в виде

$$P_{\mathsf{S}} = \Phi\big(-\sqrt{q}\tau + |c|\sqrt{q}\big) + \Phi\big(-\sqrt{q}\tau - |c|\sqrt{q}\big).$$

Подстановка $\tau\sqrt{q} = -\Phi^{-1}(P_{\mathsf{F}}/2)$ приводит к искомому результату. □

Если вероятность ложноположительного результата P_{F} достаточно мала, то выражение для вероятности успеха в теореме 4.3 хорошо аппроксимируется $P_{\mathsf{S}} \approx \Phi(\Phi^{-1}(P_{\mathsf{F}}/2) + |c|\sqrt{q})$. Следовательно, в этом случае информационная сложность имеет вид

$$q \approx \left(\frac{\Phi^{-1}(P_{\mathsf{S}}) - \Phi^{-1}(P_{\mathsf{F}}/2)}{c}\right)^2$$

в предположении, что $P_{\mathsf{S}} \geqslant P_{\mathsf{F}}/2$.

Теорему 4.3 можно использовать, даже если абсолютная величина корреляции зависит от ключа. В главе 7 показано, что критерий является оптимальным, когда от ключа зависит только знак, но не в общем случае. Чтобы проанализировать вероятность успеха в случае, когда $|c|$ зависит от ключа, выражение для P_{S} в теореме 4.3 следует усреднить по ключу. Это проиллюстрировано в следующем примере.

Пример 4.4 (возврат к примеру 2.3). В этом примере вероятность успеха различителя определяется на основе линейной аппроксимации (000000001, 000010000) для трех раундов демонстрационного шифра из раздела 1.1. В примере 2.3 было выведено следующее выражение для корреляции:

$$c = (-1)^{\kappa_1}/8\ (1 + (-1)^{\kappa_2}/2)(1 + (-1)^{\kappa_3}/2),$$

или для квадрата корреляции:

$$c^2 = \Big(\big(1 + (-1)^{\kappa_2}/2\big)\big(1 + (-1)^{\kappa_3}/2\big)\Big)^2/64.$$

Отсюда $c^2 = 1/2^{10}$ для 25 % ключей, $c^2 = 9/2^{10}$ для 50 % ключей и $c^2 = 81/2^{10}$ для 25 % ключей. На рис. 4.2 показан график зависимости вероятности успеха P_{S} от q при постоянной вероятности ложноположительного результата. Вероятность успеха равна усреднению формулы из теоремы 4.3 относительно ключа. ▷

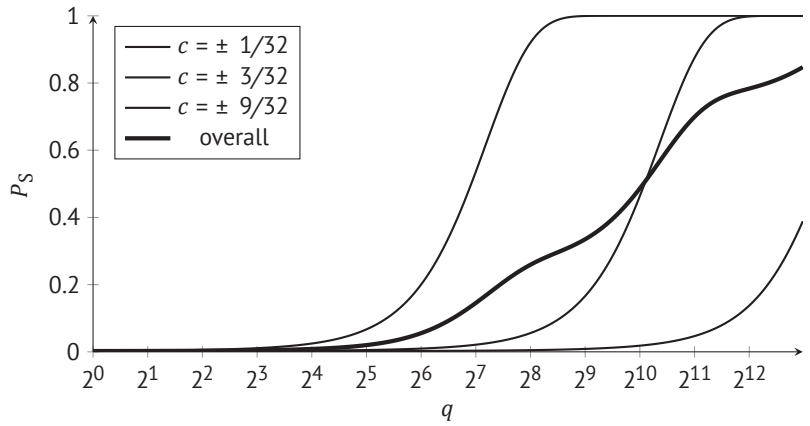


Рис. 4.2. Вероятность успеха P_S как функция от q для линейного различителя из примера 4.4 при $P_F = 0.002$. Кривая полной вероятности успеха является взвешенной суммой кривых для всех трех подмножеств ключей, т. е. overall(q) = ¼ case$_1$(q) + ½ case$_2$(q) + ¼ case$_3$(q)

4.3. Стратегии выборки

В предыдущем разделе предполагалось, что производится равномерная случайная выборка открытых текстов с возвращением. В этом случае образцы независимы и $|\{1 \leqslant i \leqslant q \mid u^{\mathrm{T}}\mathbf{x}_i = v^{\mathrm{T}}\mathbf{y}_i\}| = \sum_{i=1}^{q}(\mathbf{z}_i + 1)/2$ имеет биномиальное распределение, которое мы аппроксимировали нормальным, опираясь на центральную предельную теорему.

Если выборка открытых текстов производится без возвращения, то случайные величины $\mathbf{z}_i$ не являются независимыми: при каждой выборке значения +1 вероятность, что следующим будет выбрано значение +1, уменьшается, и наоборот. Можно показать, что в этом случае величина $|\{1 \leqslant i \leqslant q \mid u^{\mathrm{T}}\mathbf{x}_i = v^{\mathrm{T}}\mathbf{y}_i\}|$ имеет гипергеометрическое распределение, которое также можно аппроксимировать нормальным. По сравнению с нормальным распределением, аппроксимирующим биномиальное, среднее то же самое, а дисперсия оказывается меньше: она умножается на коэффициент

$$\frac{2^n - q}{2^n - 1} \approx 1 - \frac{q}{2^n},$$

который уменьшается, когда q стремится к 2^n.

Выборка с возвращением приводит к более простым формулам, а выборка без возвращения – к формулам, которые предсказывают более низкую информационную сложность. На практике, когда количество образцов q велико, могли бы возникнуть трудности с гарантией уникальности открытых текстов, потому что атакующему пришлось бы запоминать, какие значения уже встречались раньше. Эта практическая проблема не возникает, если режим работы блочного шифра гарантирует отсутствие повторений. Например, так бывает в режиме счетчика и производных от него, скажем в режиме с аутентификацией Галуа (GCM).

4.4. Восстановление ключа с использованием ранжирования ключей

Существует альтернативный подход к восстановлению ключа, который часто используется на практике. Вместо того чтобы выполнять проверку гипотез для эмпирической корреляции каждого ключа, нужно вывести список ключей-кандидатов, отсортированный по достоверности в порядке убывания (по убыванию абсолютной эмпирической корреляции). Этот подход называется *ранжирование ключей*.

Ранжирование ключей не является чем-то принципиально отличным, поскольку на практике сохраняется только часть таблицы ключей с наивысшими рангами. Как и раньше, полная атака обычно включает последний шаг, на котором угадываются и проверяются оставшиеся неизвестными биты.

Важное преимущество ранжирования ключей заключается в том, что оно терпимее к неточностям модели. Действительно, для шифров, построенных из элементов с низкой нелинейностью, таких как ARX-шифры, гипотеза рандомизации с неправильным ключом, используемая в простой модели, отклоняется от реальности.

Точно проанализировать информационную сложность ранжирования ключей труднее, чем в случае проверки гипотез. Однако если сделать дополнительные предположения, то анализ упрощается. Конкретно, ранжирование ключей можно проанализировать с применением *порядковых статистик*. Пусть $|\hat{\mathbf{c}}_1|$, $|\hat{\mathbf{c}}_2|$, ..., $|\hat{\mathbf{c}}_K|$ – абсолютные величины эмпирических корреляций K ключей. Порядковыми статистиками называются случайные величины $\mathbf{s}_1$, $\mathbf{s}_2$, ..., $\mathbf{s}_K$, полученные путем сортировки $|\hat{\mathbf{c}}_1|$, $|\hat{\mathbf{c}}_2|$, ..., $|\hat{\mathbf{c}}_K|$, так что $\mathbf{s}_1 \leqslant \mathbf{s}_2 \leqslant \ldots \leqslant \mathbf{s}_K$ с вероятностью 1. Величина $\mathbf{s}_i$ называется i-й порядковой статистикой.

Предположим, что эмпирические корреляции различных неправильных ключей независимы и одинаково распределены. Если число ключей K достаточно велико, то порядковая статистика очень похожа на квантильную функцию (обратную к функции распределения) абсолютной величины $|\hat{\mathbf{c}}_i|$ эмпирических корреляций неправильных ключей. В простой модели эта квантильная функция $p \mapsto \Phi^{-1}(p - 1)/2 / \sqrt{q}$. В частности, i-я порядковая статистика удовлетворяет приближенному равенству

$$\mathbf{s}_i \approx \frac{1}{\sqrt{q}} \Phi^{-1}\left(\frac{i-K}{2K}\right).$$

В этом равенстве знак ≈ означает, что $\mathbf{s}_i$ близка к правой части с высокой вероятностью.

Долю ключей, оставляемых в качестве кандидатов, принято записывать в виде 2^{-a}, где a называется *преимуществом восстановления ключа* (не путать с преимуществом в контексте проверки гипотез!). Ранжирование ключей приводит к успеху, если абсолютная величина эмпирической корреляции больше $K(1 - 2^{-a})$-й порядковой статистики, т. е.

$$\mathbf{s}_{\lfloor K(1-2^{-a})\rceil} \approx \frac{1}{\sqrt{q}} \Phi^{-1}\left(2^{-a-1}\right).$$

Однако это в точности совпадает с пороговым значением для проверки гипотез с $P_F = 2^{-a}$. Поэтому при тех же предположениях, что в разделе 4.2.2,

$$q \approx \left(\frac{\Phi^{-1}(P_S) - \Phi^{-1}(2^{-a-1})}{c} \right)^2.$$

Напомним (см. раздел 4.2), что величина $P_F K$ также была хорошей аппроксимацией среднего числа остающихся ключей в подходе, основанном на проверке гипотез.

4.5. Историческая справка

В своей статье о применении линейного криптоанализа к блочному шифру DES Мацуи дал оценку информационной сложности и вероятности успеха своей атаки. Селчук проанализировал процедуру ранжирования ключей более детально, опираясь на предположения простой модели из раздела 4.2. Информационная сложность выборки без возвращения была проанализирована авторами в работе Ashur, Beyne, Rijmen 2020 и независимо Блондо и Нюберг, назвавшей ее «другой известной моделью открытого текста».

Существует множество работ по статистике линейного криптоанализа, в которых рассматриваются уточнения простой модели и оптимальные методы проверки гипотез. Эти вопросы обсуждаются в главе 7.

4.6. Литература

Ashur, Tomer, Tim Beyne, and Vincent Rijmen (Apr. 2020). «Revisiting the Wrong-Key-Randomization Hypothesis». In: *Journal of Cryptology* 33.2, pp. 567–594. doi: 10.1007/s00145-020-09343-2.

Blondeau, Cerline and Kaisa Nyberg (2017). «Joint Data and Key Distribution of Simple, Multiple, and Multidimensional Linear Cryptanalysis Test Statistic and Its Impact to Data Complexity». In: *Designs, Codes and Cryptography* 82, pp. 319–349.

Matsui, Mitsuru (May 1994a). «Linear Cryptanalysis Method for DES Cipher». In: *EUROCRYPT'93*. Ed. by Tor Helleseth. Vol. 765. LNCS. Springer, Berlin, Heidelberg, pp. 386–397. doi: 10.1007/3-540-48285-7_33.

Selcëuk, Ali Aydin (Jan. 2008). «On Probability of Success in Linear and Differential Cryptanalysis». In: *Journal of Cryptology* 21.1, pp. 131–147. doi: 10.1007/s00145-007-9013-7.

4.7. Упражнения

Упражнение 4.1

Модифицируйте следствие 4.2, чтобы оно оставалось верным без предположения о «малой корреляции» из простой модели.

Упражнение 4.2

Модифицируйте теорему 4.1, следствие 4.2 и теорему 4.3 для случая, когда выборка открытых текстов производится без возвращения.

Глава 5

Методы восстановления ключа

В главе 1 было объяснено, как линейные аппроксимации можно использовать для организации атак с восстановлением ключа по алгоритму Мацуи 1 или 2. В этой главе мы более пристально рассмотрим алгоритм 2 и его улучшения. Самое важное улучшение и основная тема этой главы – «метод быстрого преобразования Фурье».

5.1. Восстановление ключа по алгоритму 2

Напомним (см. главу 1), что атака с восстановлением ключа по алгоритму Мацуи 2 разбивает шифр на внутреннюю и внешнюю части, как показано на рис. 1.4. Если у линейной аппроксимации внутренней части разреженные маски, то зачастую ее можно вычислить, зная небольшую часть открытого текста, шифртекста и раундовых ключей. Это приводит к процессу частичного шифрования и дешифрирования, показанному на рис. 5.1, где F_l и B_k имеют меньшую область определения и область значений, чем внутренняя часть E.

Прежде чем переходить к улучшениям наивного подхода, заключающегося в частичном шифровании и дешифрировании каждой пары (открытый текст, шифртекст) для каждого возможного ключа, имеет смысл систематизировать атаки с восстановлением ключа, введя терминологию для основных шагов этого процесса.

Моделирование. Первым шагом является нахождение подходящей линейной аппроксимации и определение ее корреляции (с точностью до ошибки моделирования), включая вид ее зависимости от ключа. Этот шаг был рассмотрен в главах 2 и 3.

Дистилляция. Как показано на рис. 5.1, для вычисления линейной аппроксимации внутренней части шифра необязательно знать открытый текст и шифртекст целиком. Дистилляция – это процесс извлечения релевантной информации из пар (открытый текст, шифртекст). Хотя при наивном подходе этот шаг тривиален, именно он составляет суть улучшений, обсуждаемых в разделах 5.2 и 5.3.

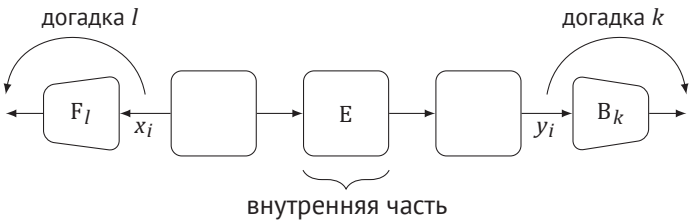


Рис. 5.1. Частичное шифрование и дешифрирование внешней части шифра

Анализ. На этом шаге данные анализируются, чтобы создать список самых вероятных (частичных) ключей. Обычно для этого нужно вычислить статистику критерия для каждого возможного ключа. Затем список статистик критерия сортируется или сравнивается с предопределенным пороговым значением. Статистические аспекты этого процесса обсуждались в главе 4.

Поиск. На последнем шаге атаки с восстановлением ключа определяется полный ключ путем исчерпывающего поиска в списке оставшихся кандидатов.

5.2. Подход Мацуи

Изложение алгоритма 2 самим Мацуи включало оптимизацию, которая опирается на более тщательный шаг дистилляции. Предположим, что имеется q пар (открытый текст, шифртекст), и обозначим i-й усеченный образец $(x_i,y_i) \in \mathbb{F}_2^s \times \mathbb{F}_2^t$. Здесь слово «усеченный» означает, что x_i и y_i включают только те биты, которые необходимы для вычисления линейной аппроксимации внутренней части шифра.

В обозначениях на рис. 5.1 существуют семейства функций $\mathsf{F}_l : \mathbb{F}_2^s \to \mathbb{F}_2$ и $\mathsf{B}_k : \mathbb{F}_2^t \to \mathbb{F}_2$ такие, что оценка корреляции линейной аппроксимации внутренней части шифра равна

$$\widehat{c}_{k,l} = \frac{1}{q}\sum_{i=1}^{q}(-1)^{\mathsf{F}_l(x_i)+\mathsf{B}_k(y_i)} = \frac{1}{q}\sum_{i=1}^{q} a_l(x_i)\, b_k(y_i), \tag{5.1}$$

где $a_l(x_i) = (-1)^{\mathsf{F}_l(x_i)}$ и $b_k(x_i) = (-1)^{\mathsf{B}_l(x_i)}$. При наивном подходе (5.1) вычисляется для каждого из K возможных значений k и для каждого из L возможных значений l, тогда полная временная сложность операций шифрования и дешифрирования равна qKL. Вышеупомянутая оптимизация улучшает этот подход, когда q больше $2^{\min\{s,t\}}$.

5.2.1. Однонаправленный случай

Сначала рассмотрим случай, когда внешняя часть состоит из одного или более раундов только в конце шифра. То есть существует лишь $L = 1$ возможных значений частичного ключа l. Чтобы подчеркнуть данный факт, опустим индексы l в формуле (5.1):

$$\widehat{c}_k = \frac{1}{q}\sum_{i=1}^{q} a(x_i)\, b_k(y_i),$$

где $a(x_i) = (-1)^{u^T x_i}$ для некоторой маски u. Правую часть этой формулы можно переписать, сгруппировав вместе члены, в которых y_i принимает одинаковое значение. Это дает формулу

$$\widehat{c_k} = \sum_{y \in \mathbb{F}_2^t} b_k(y) \frac{1}{q} \sum_{i=1}^{q} a(x_i)\, \delta^y(y_i).$$

Здесь сумму можно интерпретировать как произведение матрицы на вектор. Действительно, определим матрицу B размера $K \times 2^t$, элементы которой индексированы частичными ключами k и значениями y, и вектор w следующим образом:

$$B_{k,y} = b_k(y),$$

$$w_y = \frac{1}{q} \sum_{i=1}^{q} a(x_i)\, \delta^y(y_i).$$

При таких определениях вектор $\hat{c}$ с элементами $\hat{c}_k$ равен $\hat{c} = Bw$. Это приводит к следующим шагам дистилляции и анализа.

Дистилляция. Вычислить вектор w. Для этого требуется q вычислений a, доступов к памяти и сложений. Для сохранения вектора w требуется сохранить 2^t чисел.

Анализ. Вычислить произведение матрицы на вектор $\hat{c} = Bw$. Это произведение можно вычислить, не храня саму матрицу B. В вычислительной сложности преобладает стоимость $2^t K$ частичных дешифрирований (вычислений b_k).

Полная асимптотическая временная сложность равна $O(2^t K + q)$ по сравнению с $O(qK)$ для наивного метода.

Пример 5.1. В этом примере используется трехраундовая линейная аппроксимация из примеров 1.3 и 2.3 для организации атаки с восстановлением ключа на четыре раунда. На рис. 5.2 показана эта трехраундовая линейная аппроксимация (ненулевые маски изображены жирными линиями) и биты, участвующие в частичном дешифрировании (жирные линии). По рисунку видно, что должны быть известны три бита шифртекста и что три бита последнего раундового ключа необходимо угадать. Отсюда $K = 8$ и $t = 3$. Поскольку корреляция аппроксимации близка к $1/8$, полагаем $q = 64$.

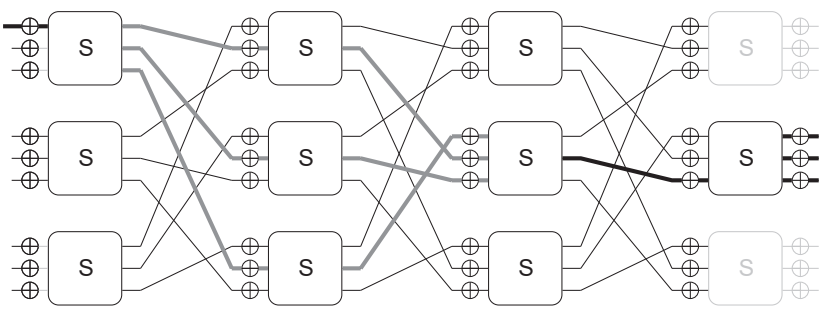


Рис. 5.2. Атака с восстановлением ключа на четыре раунда демонстрационного шифра

При использовании 64 случайных образцов, зашифрованных ключом, состоящим из одних нулей, шаг анализа включает следующее произведение матрицы на вектор (индексы, принадлежащие $\mathbb{F}_2^3$, упорядочены лексикографически):

$$\underbrace{\frac{1}{64}\begin{bmatrix}\mathbf{-24}\\ -18\\ -12\\ -14\\ 24\\ 18\\ 12\\ 14\end{bmatrix}}_{\widehat{c}} = \underbrace{\begin{bmatrix}-1 & -1 & 1 & -1 & 1 & 1 & -1 & 1\\ -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1\\ 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1\\ -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1\\ 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1\\ 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1\\ -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1\\ 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1\end{bmatrix}}_{B} \times \underbrace{\frac{1}{64}\begin{bmatrix}5\\ 6\\ 1\\ 1\\ -6\\ -4\\ -4\\ -7\end{bmatrix}}_{w}.$$

Основываясь на примере 2.3, корреляция для ключа, состоящего из одних нулей, равна $-9/32$. Поэтому эмпирическая корреляция $-24/64 = -9/32 - 3/32$ для правильного ключа – правдоподобный результат. Если знак корреляции неизвестен, то самые вероятные кандидаты – 000 и 100.

Заметим, что гипотеза рандомизации с неправильным ключом, упомянутая в главе 4, для этого примера несправедлива: корреляции для большинства неправильных ключей не близки к нулю, а примерно равны произведению $\pm 1/2$ на корреляцию для правильного ключа. Это объясняется тем, что частичное дешифрирование с неправильным ключом по сути дела добавляет в шифр один лишний (зависящий от ключа) S-блок, а самые эффективные линейные аппроксимации этого S-блока имеют корреляцию $\pm 1/2$. Большая корреляция для неправильного ключа 100 более удивительна, она объясняется в упражнении 5.1. ▷

5.2.2. Двунаправленный случай

Оптимизация, введенная в разделе 5.2.1, обобщается на случай, когда внешняя часть состоит из одного или нескольких раундов в начале u в конце шифра. В этом случае формула (5.1) переписывается следующим образом:

$$\widehat{c}_{k,l} = \frac{1}{q}\sum_{i=1}^{q} a_l(x_i)\, b_k(y_i) = \sum_{(x,y)\in\mathbb{F}_2^s\times\mathbb{F}_2^t} a_l(x)\, b_k(y)\, \frac{1}{q}\sum_{i=1}^{q} \delta^x(x_i)\delta^y(y_i).$$

Определим матрицу A размера $2^s \times L$, матрицу B размера $K \times 2^t$ и матрицу W размера $2^t \times 2^s$:

$$A_{x,l} = a_l(x),$$
$$B_{k,y} = b_k(x),$$
$$W_{y,x} = \frac{1}{q}\sum_{i=1}^{q} \delta^x(x_i)\delta^y(y_i).$$

При таких определениях матрица $\hat{c}$ размера $K \times L$ с элементами $\hat{c}_{k,l}$ равна произведению матриц BWA. Поэтому шаги дистилляции и анализа необходимо модифицировать следующим образом.

Дистилляция. Вычислить матрицу W. Для этого требуется q вычислений a и b, доступов к памяти и сложений. Для сохранения матрицы W требуется сохранить 2^{s+t} чисел.

Анализ. Вычислить произведение матриц $\hat{c} = BWA$. Его можно вычислять как $(BW)A$ или как $B(WA)$. Таким образом, стоимость вычислений равна

$$\min\{2^{s+t}LT_a + 2^tKLT_b, 2^{s+t}KT_b + 2^sKLT_a\},$$

где T_a и T_b – стоимости частичного шифрования и дешифрирования соответственно. Для этого требуется сохранить не более $KL + \max\{2^tL, 2^sK\}$ чисел.

Как правило, $K \geqslant 2^t$ и $L \geqslant 2^s$, поэтому асимптотическая вычислительная сложность составляет $O(KL2^{\min\{s,t\}} + q)$. Это улучшает наивный подход, когда $q \geqslant 2^{\min\{s,t\}}$.

В некоторых случаях развертка ключа вводит соотношения между ключами k и l. Их можно использовать для усечения матрицы $\hat{c}$. Принимая это во внимание при вычислении $\hat{c}$, можно добиться дополнительного ускорения.

5.3. Метод быстрого преобразования Фурье

Самая дорогая часть оптимизированного подхода к восстановлению ключа из раздела 5.2 – вычисление произведения матрицы на вектор Bw в однонаправленном случае и произведения матриц BWA в двунаправленном. Оказывается, что матрицы A и B часто имеют специальную структуру, благодаря которой эти операции можно ускорить.

5.3.1. Циркулянтная структура

Если раундовые ключи прибавляются к состоянию в начале и в конце шифра, то функции F_l и B_k имеют вид $\mathsf{F}_{l_1\|\,l_2}(x) = \mathsf{F}'_{l_2}(x + l_1)$ и $\mathsf{B}_{k_1\|\,k_2}(y) = \mathsf{B}'_{k2}(y + k_1)$ для некоторых функций F'_{l_2} и B'_{k_2}, индексированных ключами l_2 и k_2. Значит, существуют также a'_{l2} и b'_{k2} такие, что

$$a_{l_1\|\,l_2}(x) = a'_{l_2}(x + l_1),$$
$$b_{k_1\|\,k_2}(y) = b'_{k_2}(y + k_1).$$

Отсюда следует, что у матриц A и B своеобразная структура. Точнее, для любого l_2 обозначим A^{l_2} квадратную подматрицу A с элементами $A^{l_2}_{x,l_1} = a'_{l_2}(x + l_1)$. Для любого k_2 можно аналогично определить B^{k_2} – квадратную подматрицу B с элементами $B^{k_2}_{k_1,x} = b'_{k2}(y + k_1)$. Матрицы A^{l_2} и B^{k_2} называются *циркулянтными матрицами*.

Определение 5.1 (циркулянтная матрица). Матрица M размера $2^m \times 2^m$, индексированная элементами $\mathbb{F}_2^m$, такая что $M_{x,y} = M_{0,x+y}$ для всех $x, y \in \mathbb{F}_2^m$, называется циркулянтной.

Существует более общее определение циркулянтных матриц, в котором $\mathbb{F}_2^m$ заменено произвольной конечной группой. Возможно, вы уже знакомы с циркулянтными матрицами, индексированными циклической группой $\mathbb{Z}/N\mathbb{Z}$. Но в этом разделе мы ограничим обсуждение определением 5.1.

Пример 5.2 (циркулянтная матрица). Матрица B в примере 5.1 является циркулянтной. Например, первыми двумя ее строками будут

$$\begin{bmatrix} -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 \end{bmatrix}.$$

По сравнению с первой строкой, элементы в соседних четных и нечетных позициях (нумерация начинается с 0) переставлены. Это соответствует прибавлению 001 к индексам столбцов как элементов $\mathbb{F}_2^3$. ▷

Оказывается, что произведение циркулянтной матрицы размера $2^m \times 2^m$ на вектор можно вычислить за $O(m2^m)$ арифметических операций. При этом предполагается, что в памяти хранится одна строка матрицы. В результате анализ можно ускорить.

Однонаправленный случай. Для вычисления Bw достаточно вычислить произведение $B^{k_2}w$ для всех значений k_2. Каждое произведение матрицы на вектор требует 2^t частичных дешифрирований и $O(t2^t)$ арифметических операций. Следовательно, во временной сложности преобладает стоимость $2^tK_2 = K_1K_2$ частичных дешифрирований и $O(t2^tK_2)$ арифметических операций. Сравните со сложностью $2^tK_1K_2$ метода из раздела 5.2.

Двунаправленный случай. Произведение $B^{k_2}WA^{l_2}$ можно вычислять как $(B^{k_2}W)A^{l_2}$ или как $B^{k_2}(WA^{l_2})$. Без ограничения общности рассмотрим первый случай. Произведение $B^{k_2}W$ вычисляется путем умножения циркулянтной матрицы B^{k_2} на 2^s столбцов W. Для вычисления произведения $(B^{k_2}W)A^{l_2}$ 2^t строк $B^{k_2}W$ умножаются на циркулянтную матрицу A^{l_2}. Следовательно, в общей сложности преобладает стоимость $K_1K_2L_1L_2$ частичных шифрований *или* дешифрирований и $O(\min\{s,t\}K_1K_2L_1L_2)$ арифметических операций.

Из раздела 5.3.2 станет ясно, что многие арифметические операции можно амортизировать, когда K_2 и (или) L_2 велики, хотя общая сложность при этом не изменится.

Если $K = K_1K_2$ и $L = L_1L_2$, то общая временная сложность равна $O(KL)$ частичных шифрований и дешифрирований и $O(\min\{s,t\}KL)$ арифметических операций. Сравните со сложностью $O(2^{\min\{s,t\}}KL)$ метода из раздела 5.2.

При использовании этого метода трудно усечь матрицу $\hat{c}$, чтобы учесть потенциальные связи между K_1 и L_1, которые могли появиться в результате развертки ключа. Однако линейные связи учесть все же можно, должным образом модифицировав алгоритм умножения матриц.

5.3.2. Умножение на циркулянтные матрицы

Умножение, в котором участвует циркулянтная матрица, можно выполнить эффективно, воспользовавшись алгоритмом быстрого преобразования Фурье. Это связано с тем, что, как показывает теорема 5.3 ниже, преобразование Фурье диагонализирует циркулянтную матрицу. Преобразование Фурье (для аддитивной группы $\mathbb{F}_2^m$) определяется следующим образом.

Определение 5.2 (преобразование Фурье). Преобразованием Фурье $\mathcal{F}_m$ называется линейный оператор, который отображает вещественные векторы v, индексированные элементами $\mathbb{F}_2^m$, в вещественные векторы $\hat{v} = \mathcal{F}_m(v)$, индексированные элементами $\mathbb{F}_2^m$, следующим образом:

$$\widehat{v}_u = \sum_{x \in \mathbb{F}_2^n} (-1)^{u^\mathsf{T} x} v_x.$$

Эквивалентно, в виде матрицы относительно стандартного базиса

$$\mathcal{F}_m = \bigotimes_{i=1}^{m} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Обратная к $\mathcal{F}_m$ матрица равна $\mathcal{F}_m / 2^m$ (упражнение 5.2). Следующий результат показывает, что $\mathcal{F}_m$ диагонализирует циркулянтные матрицы.

Теорема 5.3 (диагонализация циркулянтных матриц). *Пусть M – циркулянтная матрица размера $2^m \times 2^m$ с первой строкой r. Если $\hat{r} = \mathcal{F}_m(r)$, то*

$$M = \mathcal{F}_m \begin{bmatrix} \widehat{r}_{0\cdots00} & & & \\ & \widehat{r}_{0\cdots01} & & \\ & & \ddots & \\ & & & \widehat{r}_{1\cdots11} \end{bmatrix} \mathcal{F}_m^{-1}.$$

Доказательство. Элемент (u, v) матрицы в правой части равен

$$\frac{1}{2^m} \sum_{w \in \mathbb{F}_2^m} \widehat{r}_w \, (-1)^{v^\mathsf{T} w + u^\mathsf{T} w} = \frac{1}{2^m} \sum_{w \in \mathbb{F}_2^m} \widehat{r}_w \, (-1)^{w^\mathsf{T} (u+v)}.$$

Подстановка $\widehat{r}_w = \sum_{x \in \mathbb{F}_2^m} (-1)^{w^\mathsf{T} x} \, r_x$ дает

$$\frac{1}{2^m} \sum_{w \in \mathbb{F}_2^m} \sum_{x \in \mathbb{F}_2^m} r_x \, (-1)^{w^\mathsf{T} (u+v+x)} = \sum_{x \in \mathbb{F}_2^m} r_x \, \delta^0(u + v + x) = r_{u+v}.$$

Поскольку элементы матрицы в правой части являются функциями от $u + v$, это циркулянтная матрица. Кроме того, поскольку r равна первой строке M, эта матрица равна M. □

Теорема 5.3 сразу же дает эффективный алгоритм вычисления произведения матрицы на вектор Mv. Сначала вычисляем $\mathcal{F}_m^{-1}(v)$. Для этого требуется $m2^m$ арифметических операций, как показано в упражнении 2.5. Затем вычисляем произведение диагональной матрицы и v. Для этого необходимо 2^m умножений. Наконец, вычисляем преобразование Фурье результата.

Пример 5.3. Возвращаясь к примеру 5.1, сначала вычислим преобразование Фурье первой строки M:

$$\widehat{r} = \mathcal{F}_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ -1 \\ 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 \\ -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \\ -1 \\ 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -4 \\ 4 \\ -4 \\ -4 \end{bmatrix}.$$

Затем вычислим преобразование Фурье w и умножим результат на диагональную матрицу D с вектором $\hat{r}$ на диагонали:

$$D(\mathcal{F}_3 w) = \begin{bmatrix} 0 & & & & & & & \\ & 0 & & & & & & \\ & & 0 & & & & & \\ & & & 0 & & & & \\ & & & & -4 & & & \\ & & & & & 4 & & \\ & & & & & & -4 & \\ & & & & & & & -4 \end{bmatrix} \begin{bmatrix} -8 \\ 0 \\ 10 \\ -6 \\ 34 \\ -2 \\ 8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 136 \\ -8 \\ 32 \\ 16 \end{bmatrix}.$$

Вычисление обратного преобразования Фурье дает искомый результат:

$$\widehat{c} = \mathcal{F}_3^{-1}(D\mathcal{F}_3 w) = \frac{1}{8}\mathcal{F}_3 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 136 \\ -8 \\ -32 \\ -16 \end{bmatrix} = \begin{bmatrix} -24 \\ -18 \\ -12 \\ -14 \\ 24 \\ 18 \\ 12 \\ 14 \end{bmatrix}.$$

Правая часть совпадает с результатом, приведенным в примере 5.1. ▷

5.4. Историческая справка

Выделение трех шагов в атаках с восстановлением ключа было введено Мацуи в статье по экспериментальному криптоанализу шифра DES; в ней употреблялись термины «подсчет данных» (дистилляция), «подсчет ключей» (анализ) и «исчерпывающий поиск» (поиск). Там же был впервые предложен подход, описанный в разделе 5.2. Метод быстрого преобразования Фурье из раздела 5.3 был описан в работе Коддарда, Стандаерта и Квизвотера.

5.5. Литература

Collard, Baudoin, F-X Standaert, and Jean-Jacques Quisquater (2007). «Improving the Time Complexity of Matsui's Linear Cryptanalysis». In: *Information Security and Cryptology-ICISC 2007: 10th International Conference, Seoul, Korea, November 29–30, 2007. Proceedings 10*. Springer, pp. 77–88.

Matsui, Mitsuru (Aug. 1994b). «The First Experimental Cryptanalysis of the Data Encryption Standard». In: *CRYPTO'94*. Ed. by Yvo Desmedt. Vol. 839. LNCS. Springer, Berlin, Heidelberg, pp. 1–11. doi: 10.1007/3-540-48658-5_1.

5.6. Упражнения

Упражнение 5.1

В примере 5.1 было отмечено, что неправильному ключу 100 соответствует большая эмпирическая корреляция.

1. Объясните это наблюдение и покажите, что корреляция равна в точности $^{9}/_{32}$.
2. Предположим, что правильное значение всех трех бит ключа не равно 000. Будут ли по-прежнему существовать неправильные ключи с большой корреляцией? Какие?

Упражнение 5.2

Покажите, что матрица, обратная к $\mathcal{F}_m$, равна $\mathcal{F}_m/2^m$.

Упражнение 5.3

Свертка $u \circledast v$ двух векторов u и v длины 2^m, индексированных элементами $\mathbb{F}_2^m$, определяется как

$$(u \circledast v)_y = \sum_{x \in \mathbb{F}_2^m} u_x v_{x+y}.$$

1. Пусть M – циркулянтная матрица с первой строкой u. Покажите, что $Mv = u \circledast v$.
2. Покажите, что $\mathcal{F}_m(u \circledast v) = \mathcal{F}_m(u) \odot \mathcal{F}_m(v)$, где $\odot$ – поэлементное произведение.

Глава 6

Множественный линейный криптоанализ

Если имеется более одной линейной аппроксимации, то естественно попробовать задействовать их все одновременно. Это называется множественным линейным криптоанализом. В первой части данной главы множественный линейный криптоанализ обсуждается в общих чертах. Вторая часть посвящена частному случаю, когда множество масок образует векторное пространство, – он называется многомерным линейным криптоанализом.

6.1. Множественный линейный криптоанализ

Идея множественного линейного криптоанализа заключается в том, чтобы использовать более одной линейной аппроксимации $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$.

6.1.1. Множественные линейные аппроксимации

Множественной линейной аппроксимацией функции $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ называется множество $\Lambda \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^m$ пар входных и выходных масок. Каждая пара $(u, v) \in \Lambda$ является линейной аппроксимацией F с корреляцией $C_{v,u}^{\mathsf{F}}$. *Емкость* Λ равна

$$\mathsf{Cap}(\Lambda) = \sum_{\substack{(u,v)\in\Lambda \\ (u,v)\neq(0,0)}} \left(C_{v,u}^{\mathsf{F}}\right)^2.$$

Причина введения этой величины состоит в том, что она определяет наилучшую возможную информационную сложность множественной линейной атаки, основанной на Λ. Мы обсудим это ниже и докажем в главе 7.

Для построения различителя с использованием множественной линейной аппроксимации Λ оценим каждую из корреляций $C_{v,u}^{\mathsf{F}}$, где $(u, v) \in \Lambda$. Тогда возникает задача проверки гипотез с многомерными распределениями, но ниже показано, что часто ее можно свести к одномерному случаю.

При условии что количество аппроксимаций не слишком велико по сравнению с числом образцов, используемых для оценки корреляций, многомерная центральная предельная теорема (теорема 2 из приложения А) утверждает, что совместное распределение оценок корреляций аппроксимируется многомерным нормальным распределением. По сравнению с одномерным случаем тут есть потенциальная трудность: оценки различных линейных аппроксимаций

необязательно независимы, потому что основаны на одних и тех же парах (открытый текст, шифртекст). Но многомерный нормальный случай все же поддается анализу, потому что ковариационные матрицы улавливают все зависимости. На самом деле следующий результат показывает, что нецентральные ковариации обычно пренебрежимо малы.

Теорема 6.1. *Пусть (u_1, v_1) и (u_2, v_2) – линейные аппроксимации функции* F *с эмпирическими корреляциями $\hat{\mathbf{c}}_1$ и $\hat{\mathbf{c}}_2$ соответственно. Если $\hat{\mathbf{c}}_1$ и $\hat{\mathbf{c}}_2$ оцениваются по одним и тем же q парам (открытый текст, шифртекст) с независимыми и равномерно распределенными случайными открытыми текстами, то и ковариация равна*

$$\mathsf{Cov}(\widehat{\mathbf{c}}_1, \widehat{\mathbf{c}}_2) = \left(C^{\mathsf{F}}_{v_1+v_2,\,u_1+u_2} - C^{\mathsf{F}}_{v_1,u_1} C^{\mathsf{F}}_{v_2,u_2}\right)/q.$$

Доказательство. Ковариация $\hat{\mathbf{c}}_1$ и $\hat{\mathbf{c}}_2$ равна

$$\mathsf{Cov}(\widehat{\mathbf{c}}_1, \widehat{\mathbf{c}}_2) = \mathbb{E}(\widehat{\mathbf{c}}_1\widehat{\mathbf{c}}_2) - \mathbb{E}(\widehat{\mathbf{c}}_1)\mathbb{E}(\widehat{\mathbf{c}}_2) = \mathbb{E}(\widehat{\mathbf{c}}_1\widehat{\mathbf{c}}_2) - C^{\mathsf{F}}_{v_1,u_1} C^{\mathsf{F}}_{v_2,u_2}.$$

Обозначим $(\mathbf{x}_1, \mathbf{y}_1), \ldots, (\mathbf{x}_q, \mathbf{y}_q)$ пары (открытый текст, шифртекст), использованные для вычисления $\hat{\mathbf{c}}_1$ и $\hat{\mathbf{c}}_2$. Первый член выражения выше можно раскрыть в виде

$$\begin{aligned}
\mathbb{E}(\widehat{\mathbf{c}}_1\widehat{\mathbf{c}}_2) &= \frac{1}{q^2}\sum_{i=1}^{q}\sum_{j=1}^{q}\mathbb{E}\left((-1)^{v_1^{\mathsf{T}}\mathbf{y}_i + u_1^{\mathsf{T}}\mathbf{x}_i + v_2^{\mathsf{T}}\mathbf{y}_j + u_2^{\mathsf{T}}\mathbf{x}_j}\right) \\
&= \frac{1}{q^2}\sum_{i=1}^{q}\mathbb{E}\left((-1)^{(v_1+v_2)^{\mathsf{T}}\mathbf{y}_i + (u_1+u_2)^{\mathsf{T}}\mathbf{x}_i}\right) + \frac{q(q-1)}{q^2} C^{\mathsf{F}}_{v_1,u_1} C^{\mathsf{F}}_{v_2,u_2} \\
&= \frac{1}{q} C^{\mathsf{F}}_{v_1+v_2,\,u_1+u_2} + \frac{q-1}{q} C^{\mathsf{F}}_{v_1,u_1} C^{\mathsf{F}}_{v_2,u_2}.
\end{aligned}$$

Вычитание $C^{\mathsf{F}}_{v_1,u_1} C^{\mathsf{F}}_{v_2,u_2}$ из обеих частей дает искомый результат. □

У теоремы 6.1 есть два важных следствия. Первое – то, что ковариационная матрица не полностью определяется корреляциями линейных аппроксимаций из Λ, если Λ не замкнуто относительно сложения. Второе – что ковариация оценок разных ковариаций на практике часто оказывается пренебрежимо малой. Причина в том, что если только для F не существует исключительно сильных линейных аппроксимаций, то ковариации гораздо меньше дисперсий отдельных оценок. Действительно, из теоремы 6.1 следует, что дисперсии приближенно равны $1/q$, тогда как ковариации приближенно равны c/q, где c – корреляция линейной аппроксимации F. Ковариационная матрица оценок хорошо аппроксимируется диагональной матрицей при условии, что c много меньше $1/\sqrt{|\Lambda|}$ (типичный случай) или $1/|\Lambda|$ (худший случай).

Пример 6.1. Пусть $\Lambda = \{(u_1, v_1), (u_2, v_2)\}$, где (u_1, v_1) = (000000001, 000010000) и $(u_2, v_2$ = (000000110, 000001000) – множественная линейная аппроксимация демонстрационного шифра из раздела 1.1. В силу примера 2.3, (u_1, v_1) имеет корреляцию

$$(-1)^{\kappa_1}/8\ (1 + (-1)^{\kappa_2}/2)(1 + (-1)^{\kappa_3}/2),$$

где $\kappa_1 = k_0 + k_{10} + k_{22} + k_{31} + 1$, $\kappa_2 = k_{16} + k_{21}$ и $\kappa_3 = k_{13} + k_{23}$. Аналогично корреляция (u_2, v_2) равна (проверьте!)

$$(-1)^{\lambda_1}/8\ (1 + (-1)^{\lambda_2}/2),$$

где $\lambda_1 = k_1 + k_2 + k_{16} + k_{21} + k_{30}$ и $\lambda_2 = k_{10} + k_{22} + 1$. Как следует из теоремы 6.1, ковариационная матрица $\hat{\mathbf{c}}$ зависит от корреляции линейной аппроксимации $(u_1 + u_2, v_1 + v_2)$. Анализ следов показывает, что корреляция этой аппроксимации равна

$$(-1)^{\mu_1}/8\ (1 + (-1)^{\mu_2}),$$

где $\mu_1 = k_0 + k_1 + k_2 + k_{10} + k_{22} + k_{30} + k_{31} + 1$ и $\mu_2 = \kappa_2 + \lambda_2$.

Выберем такой ключ, что $\kappa_1 = \lambda_1 = \mu_1 = \kappa_2 = \lambda_2 = 1$ и $\kappa_3 = \mu_2 = 0$. В этом случае среднее вектора эмпирических корреляций $\hat{\mathbf{c}}$ равно

$$\mathbb{E}(\hat{\mathbf{c}}) = -\begin{bmatrix} 3/32 \\ 1/16 \end{bmatrix}.$$

Ковариационная матрица $\mathbb{E}((\hat{\mathbf{c}} - \mathbb{E}(\hat{\mathbf{c}}))(\hat{\mathbf{c}} - \mathbb{E}(\hat{\mathbf{c}}))^{\mathrm{T}})$ равна

$$\frac{1}{q}\begin{bmatrix} 1 & -1/4 \\ -1/4 & 1 \end{bmatrix} - \frac{1}{q}\begin{bmatrix} 9/1024 & 3/512 \\ 3/512 & 1/256 \end{bmatrix}.$$

Вторым членом в выражении выше можно пренебречь.

Некоторые выборочные оценки корреляций показаны на рис. 6.1. В этом случае ковариацией пренебречь нельзя, потому что вектор $(u_1 + u_2, v_1 + v_2)$ был выбран так, что его абсолютная корреляция очень велика. Линии постоянства плотности вероятности (например, эллипс на рис. 6.1) обычно более близки к окружности. ▷

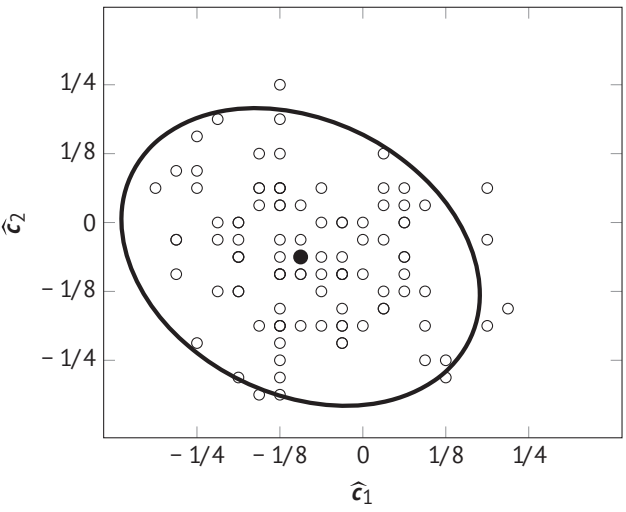


Рис. 6.1. Оценки корреляций с $q = 64$ образцами (100 образцами)

6.1.2. Различители

Пусть $\hat{\mathbf{c}}_1, \ldots, \hat{\mathbf{c}}_{|\Lambda|}$ – оценки корреляций линейных аппроксимаций для множественной линейной аппроксимации Λ функции F. Предположим, что Λ не содержит тривиальной линейной аппроксимации (0, 0). Для статистического анализа используется простая модель из главы 4 вкупе с некоторыми дополнительными предположениями, которые мы объясним ниже.

Известные корреляции. Предположим, что все корреляции $C^{\mathsf{F}}_{v,u}$, где $(u, v) \in \Lambda$ известны и что все нецентральные ковариации пренебрежимо малы. В принципе, построение различителя с использованием Λ – задача многомерной статистики. Однако мы можем свести ее к одномерной, взяв линейную комбинацию оценок $\hat{\mathbf{c}}_1, \ldots, \hat{\mathbf{c}}_{|\Lambda|}$ в качестве статистики критерия:

$$\mathbf{t}_\Lambda = \sum_{i=1}^{|\Lambda|} w_i \, \hat{\mathbf{c}}_i.$$

Чтобы дисперсия $\mathbf{t}_\Lambda$ оставалась постоянной (равной $1/q$ с точностью до небольшой погрешности), веса $w_1, \ldots, w_{|\Lambda|}$ должны удовлетворять соотношению $\sum_{i=1}^{|\Lambda|} w^2_i = 1$. Если образцы случайны и равномерно распределены, то среднее $\mathbf{t}_\Lambda$ равно нулю – при условии что $(0, 0) \notin \Lambda$. Однако если образцы выбраны из шифра, то среднее $\mathbf{t}_\Lambda$ равно

$$\mathbb{E}(\mathbf{t}_\Lambda) = \sum_{i=1}^{|\Lambda|} w_i \, C^{\mathsf{F}}_{v_i, u_i}.$$

Теорема 4.1 показывает, что информационная сложность различителя, основанная на статистике критерия $\mathbf{t}_\Lambda$, обратно пропорциональна квадрату $\mathbb{E}(\mathbf{t}_\Lambda)$. Поэтому имеет смысл максимизировать $\mathbb{E}(\mathbf{t}_\Lambda)$, оставляя дисперсию постоянной. В силу упражнения 6.1 эта цель достигается, если выбрать

$$w_i = \frac{C^{\mathsf{F}}_{v_i, u_i}}{\sqrt{\sum_{i=1}^{|\Lambda|} \left(C^{\mathsf{F}}_{v_i, u_i}\right)^2}}.$$

В таком случае среднее равно $\sqrt{\mathsf{Cap}(\Lambda)}$. Поэтому, в силу теоремы 4.1, информационная сложность пропорциональна $1/\mathsf{Cap}(\Lambda)$. Точнее, в простой модели с пренебрежимо малыми нецентральными ковариациями информационная сложность q равна

$$q = \frac{\left(\Phi^{-1}(P_{\mathsf{S}}) - \Phi^{-1}(P_{\mathsf{F}})\right)^2}{\mathsf{Cap}(\Lambda)},$$

где P_{S} – вероятность успеха, а $P_{\mathsf{F}} \leqslant P_{\mathsf{S}}$ – вероятность ложноположительного результата. В главе 7 показано, что это значение, по существу, оптимально.

Пример 6.2. Рассмотрим множественную линейную аппроксимацию из примера 6.1. Для того же ключа, что и раньше, емкость равна $^{13}/_{1024}$, и статистика критерия $\mathbf{t}_\Lambda$ равна

$$\mathbf{t}_\Lambda = -\frac{3}{\sqrt{13}}\widehat{\mathbf{c}}_1 - \frac{2}{\sqrt{13}}\widehat{\mathbf{c}}_2.$$

Гистограмма распределения статистики критерия показана на рис. 6.2. ▷

Неизвестные корреляции. Корреляции линейных аппроксимаций обычно зависят от ключа, поэтому использовать описанную выше стратегию без угадывания битов ключа невозможно. Систематический способ решения этой проблемы представлен в главе 7. Неоптимальный общий метод и лучший метод для случая, когда знаки корреляций неизвестны, обсуждаются ниже.

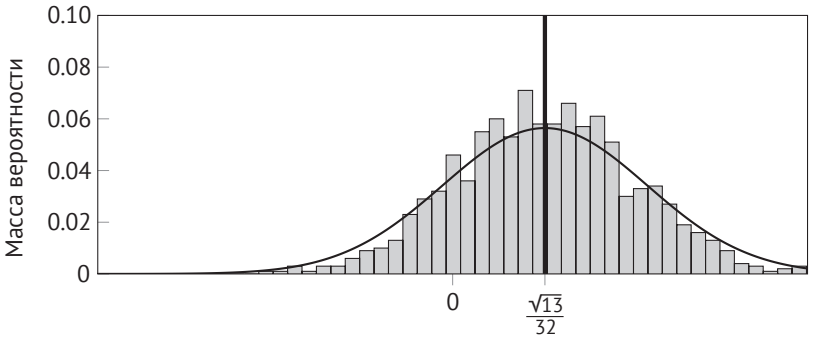


Рис. 6.2. Гистограмма статистики критерия $\mathbf{t}_\Lambda$ для 1000 экспериментов

Простой практический критерий основан на линейной комбинации квадратов оценок корреляций:

$$\mathbf{t}_\Lambda = \sum_{i=1}^{|\Lambda|} w_i \left(\widehat{\mathbf{c}}_i\right)^2.$$

Статистика критерия $\mathbf{t}_\Lambda$ не является нормально распределенной, даже если все оценки $\hat{\mathbf{c}}_1, \ldots, \hat{\mathbf{c}}_{|\Lambda|}$ являются таковыми. Тем не менее если значение $|\Lambda|$ велико, то нормальное распределение будет хорошим приближением. Поэтому для оценки информационной сложности при этом предположении достаточно определить среднее и дисперсию $\mathbf{t}_\Lambda$. Следующая статистическая лемма дает нужный нам результат.

Лемма 6.2. *Пусть* $\mathbf{x}_1, \ldots, \mathbf{x}_l$ *– попарно некоррелированные случайные величины со средними* $\mu_1, \ldots, \mu_l$ *и дисперсией* σ^2. *Среднее* $\sum_{i=1}^{l} w_i \mathbf{x}^2$ *равно* $\sum_{i=1}^{l} w_i(\sigma^2 + \mu_i^2)$. *Кроме того, если* $\mathbf{x}_1, \ldots, \mathbf{x}_l$ *нормально распределены, то дисперсия равна* $2\sigma_l^2 \sum_{i=1}^{l} w_i^2(\sigma^2 + 2\mu_i^2)$.

Доказательство. Результат, касающийся среднего, следует из того факта, что среднее $\mathbf{x}_i^2$ равно $\sigma^2 + \mu_i^2$ и что $\mathbb{E}(\mathbf{x}_i\mathbf{x}_j) = 0$ для $i \neq j$. Если $\mathbf{x}_1, \ldots, \mathbf{x}_{n'}$ нормально распределены и попарно некоррелированы, то они независимы, поэтому их квадраты также попарно некоррелированы. Результат, касающийся дисперсии, следует из того, что

$$\begin{aligned}\mathbb{V}(\mathbf{x}_i^2) &= \mathbb{E}\big((\mathbf{x}_i-\mu_i)^4\big) + 6\mu_i^2\sigma^2 + \mu_i^4 - \big(\mu_i^2+\sigma^2\big)^2\\ &= \mathbb{E}\big((\mathbf{x}_i-\mu_i)^4\big) + 4\mu_i^2\sigma^2 - \sigma^4\\ &= 2\sigma^4 + 4\mu_i^2\sigma^2.\end{aligned}$$

Равенство в первой строке – результат несложного вычисления, см. упражнение 6.2. На последнем шаге используется тот факт, что четвертый момент нормального распределения $\mathcal{N}(0,1)$ равен трем. Это можно доказать с помощью интегрирования по частям:

$$\int_{-\infty}^{\infty} x^4 e^{-x^2/2}\,\mathrm{d}x = 2\int_0^{\infty} x^3\left(-x\,e^{-x^2/2}\right)\mathrm{d}x = 3\int_{-\infty}^{\infty} x^2 e^{-x^2/2}\,\mathrm{d}x,$$

где нормировочный коэффициент $1/\sqrt{2\pi}$ опущен. □

Если образцы равномерно распределены, то, по лемме 6.2 с $\mu_1 = \mu_2 = \ldots = 0$ и $\sigma^2 = 1$, среднее $\mathbf{t}_\Lambda$ равно $\sum_{i=1}^{|\Lambda|} w_i/q$ и дисперсия равна $2/q^2$, если $\sum_{i=1}^{|\Lambda|} w_i^2 = 1$. Если образцы выбраны из шифра, то среднее равно (в силу леммы 6.2 с $\mu_i = C^{\mathsf{F}}_{v_i,u_i}$ и $\sigma^2 \approx 1/q$)

$$\mathbb{E}(\mathbf{t}_\Lambda) = \frac{1}{q}\sum_{i=1}^{|\Lambda|} w_i + \sum_{i=1}^{|\Lambda|} w_i\big(C^{\mathsf{F}}_{v_i,u_i}\big)^2.$$

Выражение для дисперсии длиннее, но оно близко к $2/q^2$ при условии, что q мало по сравнению с $1/(C^{\mathsf{F}}_{v,u})^2$ для всех $(u, v) \in \Lambda$. Это предположение разумно, т. к. в противном случае одной аппроксимации уже было бы достаточно. Сдвиг $\mathbf{t}_\Lambda$ на $\sum_{i=1}^{|\Lambda|} w_i/q$ показывает, что задача проверки гипотез сводится к различению распределений $\mathcal{N}(\sum_{i=1}^{|\Lambda|} w_i(C^{\mathsf{F}}_{v_i,u_i})^2,\ 2/q^2)$ и $\mathcal{N}(0,\ 2/q^2)$. По теореме 4.1 с q^2 вместо q имеем

$$q = \sqrt{2}\,\frac{\Phi^{-1}(P_{\mathsf{S}}) - \Phi^{-1}(P_{\mathsf{F}})}{\sum_{i=1}^{|\Lambda|} w_i\big(C^{\mathsf{F}}_{v_i,u_i}\big)^2}, \tag{6.1}$$

где P_{S} – вероятность успеха, а $P_{\mathsf{F}} \leqslant P_{\mathsf{S}}$ – вероятность ложноположительного результата.

Эта формула опирается на предположения простой модели из раздела 4.2, а также на вышеупомянутые предположения: нецентральные ковариации должны быть пренебрежимо малы, $|\Lambda|$ должно быть велико, а q мало по сравнению с $1/(C^{\mathsf{F}}_{v,u})^2$ для всех $(u, v) \in \Lambda$. Эти дополнительные предположения зачастую разумны. Однако если абсолютные корреляции близки к $2^{-n/2}$, то простая модель становится ненадежной. Данный вопрос мы еще обсудим в разделе 7.3.3. Если о корреляциях вообще ничего не известно, то лучшее, что можно сделать, – выбрать равные веса $w_1 = \ldots = w_{|\Lambda|} = 1/\sqrt{|\Lambda|}$. В силу (6.1), это приводит к информационной сложности

$$q = \sqrt{2|\Lambda|}\,\frac{\Phi^{-1}(P_{\mathsf{S}}) - \Phi^{-1}(P_{\mathsf{F}})}{\mathsf{Cap}(\Lambda)}.$$

Данный подход неоптимален, потому что почти всегда имеется зависящее от ключа выражение для корреляций, которое помогает составить какое-то представление об их значениях.

Простейший пример – когда известны абсолютные величины корреляций, но неизвестны их знаки. В таком случае следует использовать веса, пропорциональные $(C^{\mathsf{F}}_{v_i,u_i})^2$. В силу (6.1), информационная сложность пропорциональна

$$\frac{1}{\sqrt{\sum_{i=1}^{|\Lambda|}\left(C^{\mathsf{F}}_{v_i,u_i}\right)^4}} \leq \frac{\sqrt{|\Lambda|}}{\mathsf{Cap}(\Lambda)}.$$

Верхняя граница правой части достигается, когда абсолютные корреляции всех линейных аппроксимаций в Λ равны. В главе 7 будет показано, что в общем случае этот критерий все равно не оптимален. Когда знаки корреляций различных аппроксимаций зависят от одних и тех же битов ключа, информационную сложность часто можно уменьшить.

Даже если квадраты корреляций также зависят от ключа, приведенный выше критерий все равно можно использовать (применяя какую-то оценку квадратов корреляций, например среднее, чтобы определить $w_1, \ldots, w_{|\Lambda|}$), но в этом случае оценивание информационной сложности технически труднее, потому что формулу вероятности успеха следует усреднять по ключу, как объяснено в разделе 4.2.2. В главе 7 будет показано, что в некоторых случаях можно добиться большего.

6.2. Многомерный линейный криптоанализ

Многомерная линейная аппроксимация – это множественная линейная аппроксимация Λ такая, что Λ является векторным пространством над полем $\mathbb{F}_2$. Поскольку многомерные линейные аппроксимации – частный случай множественных линейных аппроксимаций, их можно использовать для построения различителей точно так же, как было описано в разделе 6.1.2. Однако тот факт, что Λ является векторным пространством, приводит к интересному альтернативному описанию этих различителей.

6.2.1. Многомерные линейные аппроксимации

Первый намек на то, что многомерные линейные аппроксимации являются чем-то особенным, дает теорема 6.1: если Λ – векторное пространство, то ковариационная матрица оценок корреляций полностью определяется корреляциями линейных аппроксимаций из Λ. Тому есть веская причина: многомерная линейная аппроксимация эквивалентна линейной проекции пар (открытый текст, шифртекст). Чтобы уточнить это заявление, нам понадобятся некоторые понятия и факты из линейной алгебры.

Пусть U – векторное пространство над $\mathbb{F}_2$. Для любого подпространства $V \subseteq U$ факторпространство $U/V = \{x + V \mid x \in U\}$ является векторным пространством размерности $\dim U - \dim V$. Проекция $\pi_V : x \mapsto x + V$ является линейным отображением U в U/V. Для любых $x, y \in U$ эти понятия связаны следующим образом:

$$x \equiv y \bmod V \Leftrightarrow \pi_V(x) = \pi_V(y) \Leftrightarrow x - y \in V.$$

Предположим, что U снабжено симметричной билинейной формой $(x, y) \mapsto x \cdot y$, или «скалярным произведением». Это бинарная операция, такая что $x \cdot y = y \cdot x$, $0 \cdot x = 0$ и $(x + y) \cdot z = x \cdot z + y \cdot z$. *Ортогональным дополнением* подпространства $V \subseteq U$ называется векторное пространство

$$V^{\perp} = \{ x \in U \mid x \cdot y = 0 \text{ для всех } y \in V\}.$$

Хотя $\dim V^{\perp} = \dim U - \dim V$, подпространство $V^{\perp}$ не является дополнением V в алгебраическим смысле: может случиться, что $V \cap V^{\perp} \neq \{0\}$.

Пример 6.3 Пусть $U = \mathbb{F}_2^n \times \mathbb{F}_2^m$ и $V = \Lambda$. Скалярное произведение (u, v) и $(x, y) \in U$ равно $(u, v) \cdot (x, y) = u^{\mathsf{T}}x + v^{\mathsf{T}}y$. Отсюда

$$\Lambda^{\perp} = \{(x, y) \in \mathbb{F}_2^n \times \mathbb{F}_2^m \mid u^{\mathsf{T}}x + v^{\mathsf{T}}y = 0 \text{ для любых } (u, v) \in \Lambda\}.$$

Ниже факторпространство $(\mathbb{F}_2^n \times \mathbb{F}_2^m)/\Lambda^{\perp}$ будет играть важную роль. Это векторное пространство той же размерности, что Λ. ▷

Используя введенные выше понятия, мы можем сформулировать теорему 6.3. В анализе Фурье этот результат известен как *формула суммирования Пуассона*. Эта связь будет объяснена в главах 10 и 11.

Теорема 6.3. *Пусть* $\mathbf{z}$ *– случайная величина в векторном пространстве* U *со скалярным произведением* $(x, y) \mapsto x \cdot y$, *и пусть* V *– подпространство* U. *Для любых* $t \in U$

$$\Pr\left[\mathbf{z} \equiv t \bmod V^{\perp}\right] = \frac{1}{|V|} \sum_{v \in V} (-1)^{v \cdot t} c_{v \cdot \mathbf{z}},$$

где $c_{v \cdot \mathbf{z}}$ *обозначает корреляцию случайной величины* $v \cdot \mathbf{z}$.

Доказательство. Положим $p_{\mathbf{z}}(t) = \Pr[\mathbf{z} = t]$. Напомним (см. раздел 2.1), что

$$c_{v \cdot \mathbf{z}} = \sum_{z \in U} (-1)^{v \cdot z} p_{\mathbf{z}}(z).$$

Подставляя это в сумму, получаем

$$\sum_{v \in V} (-1)^{v \cdot z} c_{v \cdot \mathbf{z}} = \sum_{v \in V} \sum_{z \in U} (-1)^{v \cdot (z+t)} p_{\mathbf{z}}(z) = \sum_{z \in U} p_{\mathbf{z}}(z) \sum_{v \in V} (-1)^{v \cdot (z+t)}.$$

Внутреннюю сумму можно вычислить, применяя тот же подход, что в доказательстве теоремы 2.4:

$$\sum_{v \in V} (-1)^{v \cdot (z+t)} = \begin{cases} |V|, & \text{если } \ z + t \in V^{\perp}, \\ 0 & \text{в противном случае.} \end{cases}$$

Теперь результат следует из того, что $\Pr[\mathbf{z} \equiv t \bmod V^{\perp}] = \sum_{z \in t + V^{\perp}} p_{\mathbf{z}}(z)$. □

Применение теоремы 6.3 к случаю многомерной линейной аппроксимации дает следующий результат.

Следствие 6.4. *Пусть* Λ *– многомерная линейная аппроксимация функции* F: $\mathbb{F}_2^n \to \mathbb{F}_2^m$. *Если* $\mathbf{x}$ *– равномерная случайная величина на* $\mathbb{F}_2^n$, *то*

$$\Pr\left[(\mathbf{x}, \mathsf{F}(\mathbf{x})) \equiv (s,t) \bmod \Lambda^{\perp}\right] = \frac{1}{|\Lambda|} \sum_{(u,v)\in\Lambda} (-1)^{u^{\mathsf{T}}s + v^{\mathsf{T}}t} C^{\mathsf{F}}_{v,u}$$

для всех $s \in \mathbb{F}_2^n$ *и* $t \in \mathbb{F}_2^m$. *Это отношение обратимо.*

Доказательство. Воспользуемся теоремой 6.3 с $U = \mathbb{F}_2^n \times \mathbb{F}_2^m$ и $V = \Lambda$, как в примере 6.3. Случайная величина $\mathbf{z}$ равна $(\mathbf{x}, \mathsf{F}(\mathbf{x}))$. Отношение в теореме 6.3 обратимо; нахождение обратного отношения составляет предмет упражнения 6.3. □

Следствие 6.4 показывает, что корреляции линейных аппроксимаций из Λ определяют распределение вероятностей $\pi_{\Lambda^{\perp}}((\mathbf{x}, \mathsf{F}(\mathbf{x}))$. Это линейная проекция входных и выходных битов. В главе 11 показано, что отношение между корреляциями и распределением вероятностей $\pi_{\Lambda^{\perp}}((\mathbf{x}, \mathsf{F}(\mathbf{x}))$ описывается преобразованием Фурье, и объясняется, почему это так.

Пример 6.4. Пусть $\Lambda = \{(0, 0), (u_1, v_1), (u_2, v_2), (u_1 + u_2, v_1 + v_2)\}$, где $(u_1, v_1) = (000000001, 000010000)$ и $(u_2, v_2) = (000000110, 000100000)$. Ортогональное дополнение $\Lambda^{\perp}$ состоит из всех пар (x, y) таких, что $x_0 + y_4 = 0$ и $x_1 + x_2 + y_5 = 0$, где $(x_8, \ldots, x_0)$ и $(y_8, \ldots, y_0)$ – элементы x и y соответственно. Следовательно, возможным базисом $(\mathbb{F}_2^9 \times \mathbb{F}_2^9)/\Lambda^{\perp} \sim= \mathbb{F}_2^2$ является

$$(000000000,000010000) + \Lambda^{\perp}, (000000000,000100000) + \Lambda^{\perp}.$$

В этом базисе координаты проекции $(x, y) \bmod \Lambda^{\perp}$ точки (x, y) равны

$$(x_0 + y_4, x_1 + x_2 + y_5).$$

В силу следствия 6.4, вероятность того, что $(\mathbf{x}, \mathsf{F}(\mathbf{x})) \equiv (0, 0) \bmod \Lambda^{\perp}$, равна $1/4 (1 - 3/32 - 1/16 - 1/4) = 19/128$. ▷

Из следствия 6.4 следует, что линейную аппроксимацию можно выразить по-другому. Правая часть в следующей теореме называется *квадратичным евклидовым расхождением*. Доказательство этого результата составляет предмет упражнения 6.4.

Следствие 6.5 (*квадратичное евклидово расхождение*). *Пусть* Λ *– многомерная линейная аппроксимация функции* F: $\mathbb{F}_2^n \to \mathbb{F}_2^m$. *Если* $\mathbf{x}$ *– случайная равномерно распределенная величина на* $\mathbb{F}_2^n$, *то*

$$\mathsf{Cap}(\Lambda) = |\Lambda| \sum_{z} \left(\Pr\left[(\mathbf{x}, \mathsf{F}(\mathbf{x})) \equiv z \bmod \Lambda^{\perp}\right] - \frac{1}{|\Lambda|}\right)^2,$$

где суммирование производится по всем $z \in (\mathbb{F}_2^n \times \mathbb{F}_2^m)/\Lambda^{\perp}$.

6.2.2. Различители

Поскольку всякая многомерная линейная аппроксимация является множественной линейной аппроксимацией, различители можно получить так же, как описано в разделе 6.1.2. Однако в свете следствия 6.4 существует и другой подход.

Вместо того чтобы оценивать корреляции линейных аппроксимаций из Λ, можно оценить распределение вероятностей $\pi_{\Lambda^\perp}((\mathbf{x}, \mathsf{F}(\mathbf{x})))$ для случайной равномерно распределенной величины $\mathbf{x}$.

Известные корреляции. Если все корреляции известны, то известно и распределение вероятностей $\pi_{\Lambda^\perp}((\mathbf{x}, \mathsf{F}(\mathbf{x})))$. В этом случае можно воспользоваться статистикой критерия

$$\mathbf{t}_\Lambda = \sum_{i=1}^{|\Lambda|} w_i \, (\widehat{\mathbf{p}}_i - p_i),$$

где $p_1, \ldots, p_{|\Lambda|}$ – вероятности $|\Lambda|$ значений, принадлежащих $(\mathbb{F}_2^n \times \mathbb{F}_2^m)/\Lambda^\perp$, а $\mathbf{p}_1, \ldots, \mathbf{p}_{|\Lambda|}$ – их оценки. В упражнении 6.5 вам будет предложено показать, что выбор оптимальных весов $w_1, \ldots, w_{|\Lambda|}$ приводит к различителю, информационная сложность которого обратно пропорциональна квадратичному евклидову расхождению. В силу следствия 6.5, квадратичное евклидово расхождение равно $\mathsf{Cap}(\Lambda)$, поэтому информационная сложность такая же, как в разделе 6.1.2.

Неизвестные корреляции. Если корреляции неизвестны, то популярным подходом является *критерий Пирсона* χ^2. Он основан на статистике критерия

$$\mathbf{t}_\Lambda = \sum_{i=1}^{|\Lambda|} \frac{\left(\widehat{\mathbf{p}}_i - 1/|\Lambda|\right)^2}{1/|\Lambda|}.$$

В упражнении 6.6 вам будет предложено показать, что информационная сложность этого критерия пропорциональна $\sqrt{|\Lambda|}/\,\mathsf{Cap}(\Lambda)$. Это то же самое, что для критерия с равными весами из раздела 6.1.2, но хуже, чем критерий для неизвестных корреляций с известной абсолютной величиной.

6.2.3. Атаки с выбранным открытым текстом

До сих пор предполагалось, что входом примитива является случайная равномерно распределенная величина на $\mathbb{F}_2^n$. Из следствия 6.4 вытекает также, что многомерные линейные аппроксимации могут что-то сказать о выходе, когда входом является случайная равномерно распределенная величина на аффинном подпространстве $\mathbb{F}_2^n$. В некоторых случаях это наблюдение полезно, чтобы уменьшить информационную сложность.

Если $\Lambda = \Lambda_{\text{in}} \oplus \Lambda_{\text{out}}$, то следствие 6.4 принимает такой вид.

Следствие 6.6. *Пусть Λ – многомерная линейная аппроксимация функции $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$. Предположим, что $\Lambda = \Lambda_{\text{in}} \oplus \Lambda_{\text{out}}$, где $\Lambda_{\text{in}} \subseteq \mathbb{F}_2^n$ и $\Lambda_{\text{out}} \subseteq \mathbb{F}_2^m$. Если $\mathbf{x}$ – случайная равномерно распределенная величина на $s + \Lambda_{\text{in}}^\perp$, то*

$$\Pr\left[\mathsf{F}(\mathbf{x}) \equiv t \bmod \Lambda_{\text{out}}^\perp\right] = \frac{1}{|\Lambda_{\text{out}}|} \sum_{(u,v)\in\Lambda} (-1)^{u^\intercal s + v^\intercal t} \, C^{\mathsf{F}}_{v,u}$$

для всех $t \in \mathbb{F}_2^m$.

Доказательство. Если $\Lambda = \Lambda_{\text{in}} \oplus \Lambda_{\text{out}}$, то также $\Lambda^\perp = \Lambda_{\text{in}}^\perp \oplus \Lambda_{\text{out}}^\perp$. Следовательно, если $\mathbf{x}'$ – случайная равномерно распределенная величина на $\mathbb{F}_2^n$, $\mathbf{x}$ – случайная равномерно распределенная величина на $s + \Lambda_{\text{in}}^\perp$, то

$$\Pr\left[(\mathbf{x}', \mathsf{F}(\mathbf{x}')) \equiv (s,t) \bmod \Lambda^{\perp}\right] = \Pr\left[\mathbf{x}' \equiv s \bmod \Lambda_{\mathsf{in}}^{\perp} \wedge \mathsf{F}(\mathbf{x}') \equiv t \bmod \Lambda_{\mathsf{out}}^{\perp}\right]$$
$$= \frac{\Pr\left[\mathsf{F}(\mathbf{x}) \equiv t \bmod \Lambda_{\mathsf{out}}^{\perp}\right]}{|\Lambda_{\mathsf{in}}|},$$

где на втором шаге было использовано равенство $|\Lambda_{\mathrm{in}}^{\perp}|/2^n = 1/|\Lambda_{\mathrm{in}}|$. Наконец, из следствия 6.4 вытекает, что

$$\frac{\Pr\left[\mathsf{F}(\mathbf{x}) \equiv t \bmod \Lambda_{\mathsf{out}}^{\perp}\right]}{|\Lambda_{\mathsf{in}}|} = \frac{1}{|\Lambda|} \sum_{(u,v)\in\Lambda} (-1)^{u^{\mathsf{T}}s + v^{\mathsf{T}}t} C_{v,u}^{\mathsf{F}}.$$

Умножение на $|\Lambda_{\mathrm{in}}|$ дает искомый результат, потому $|\Lambda_{\mathrm{in}}||\Lambda_{\mathrm{out}}| = |\Lambda|$. □

В силу следствия 6.6, можно настроить различитель с выбранным открытым текстом, выбирая открытые тексты из аффинного подпространства $\mathbb{F}_2^n$ и оценивая распределение проекции шифртекстов. Такие атаки также называются *статистическими атаками с насыщением*. Эта терминология будет объяснена в главе 9.

Если корреляции (а значит, и распределения вероятностей) известны, то информационная сложность статистического критерия из раздела 6.2.2 обратно пропорциональна квадратичному евклидову расхождению, которое равно

$$|\Lambda_{\mathsf{out}}| \sum_{t} \left(\Pr\left[\mathsf{F}(\mathbf{x}) \equiv t \bmod \Lambda_{\mathsf{out}}^{\perp}\right] - \frac{1}{|\Lambda_{\mathsf{out}}|} \right)^2,$$

где суммирование производится по всем $t \in \mathbb{F}_2^m/\Lambda_{\mathrm{out}}^{\perp}$. Хотя эта величина зависит от выбора смежного класса $s+\Lambda_{\mathrm{in}}^{\perp}$, ее среднее равно $\mathsf{Cap}(\Lambda)$ при равномерном случайном выборе s. Следовательно, от того, что все корреляции известны, информационная сложность, как правило, не улучшается[1].

Однако когда корреляции неизвестны, информационная сложность статистического критерия из раздела 6.2.2 равна $\sqrt{|\Lambda_{\mathsf{out}}|}/\,\mathsf{Cap}(\Lambda)$, а не $\sqrt{|\Lambda|}/\,\mathsf{Cap}(\Lambda)$. Стало быть, использование выбранных открытых текстов приводит к информационной сложности, меньшей в $\sqrt{|\Lambda|/|\Lambda_{\mathsf{out}}|} = \sqrt{|\Lambda_{\mathsf{in}}|}$ раз.

6.3. Заключительные замечания

Прежде чем завершить обсуждение множественного линейного криптоанализа, уместно будет сделать несколько замечаний о вопросах, которые мы опустили.

6.3.1. Восстановление ключа

Простой способ применить методы восстановления ключа из главы 5 к множественному линейному криптоанализу – повторить обычные алгоритмы $|\Lambda|$ раз: по одному для каждой аппроксимации. Недостаток такого подхода заключается в том, что временная сложность шага анализа оказывается в $|\Lambda|$ раз больше. Следовательно, хотя множественный линейный криптоанализ уменьшает информационную сложность, он, возможно, никак не улучшит временную сложность атаки с восстановлением ключа.

[1] Этот вывод не учитывает выигрыш при использовании выборки без возвращения.

Однако часто можно достичь большего. Например, легко видеть, что улучшения возможны, когда некоторые аппроксимации используют общую входную (или выходную) маску. Существует систематический подход к этой проблеме, но его описание требует лучшего понимания множественного линейного криптоанализа. Мы отложим это до главы 11.

Наконец, двунаправленные алгоритмы восстановления ключа из главы 5 невозможно сочетать с использованием выбранных открытых текстов (как, собственно, и выбранных шифртекстов), как в разделе 6.2.3. Для гарантии правильной структуры открытых текстов в общем случае необходимы дополнительные образцы. Однако существуют интересные исключения, например когда аффинное подпространство открытых текстов используется совместно с частичным шифрованием уровня сложения с ключом.

6.3.2. Нахождение подходящих линейных аппроксимаций

В этой главе (и в разделе 6.1 в частности) много внимания было уделено статистическим аспектам множественного линейного криптоанализа. Иными словами, мы обсуждали, как использовать несколько линейных аппроксимаций, а не как их находить.

В принципе, методов из глав 2 и 3 достаточно для нахождения подходящих линейных аппроксимаций. Однако самые мощные множественные линейные атаки конструируют множественные линейные аппроксимации раунд за раундом. К этому вопросу мы также вернемся после прочтения главы 11.

6.4. Историческая справка

Множественный линейный криптоанализ был предложен Калиски и Робшоу. Бирюков, Де Канниере и Квизквотер проанализировали статистику множественного линейного криптоанализа в случае, когда корреляции известны, а от ключа зависит только их знак. Анализ, применимый к случаю неизвестных корреляций, можно найти в работе Блондо и Нюберг.

Многомерный линейный криптоанализ впервые был предложен в работе Эрмелин, Чо и Нюберг. Как объясняется в разделе 6.2, многомерный линейный криптоанализ в первую очередь интересуется связью с распределениями линейных проекций открытых и шифртекстов в постановке с известным (следствие 6.4) и выбранным (следствие 6.6) открытым текстом. Использование критерия Пирсона χ^2 в криптоанализе предшествует многомерным линейным аппроксимациям и впервые было предложено Воденэ.

6.5. Литература

Biryukov, Alex, Christophe De Canni`ere, and MichaËel Quisquater (2004). «On Multiple Linear Approximations». In: *Advances in Cryptology – CRYPTO 2004, 24th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15–19, 2004, Proceedings*. Ed. by Matthew K. Franklin. Vol. 3152. LNCS. Springer, Berlin, Heidelberg, pp. 1–22. doi: 10.1007/978-3-540-28628-8_1. url: https://doi.org/10.1007/978-3-540-28628-8%5C_1.

Blondeau, Cerline and Kaisa Nyberg (2017). «Joint Data and Key Distribution of Simple, Multiple, and Multidimensional Linear Cryptanalysis Test Statistic and Its Impact to Data Complexity». In: *Designs, Codes and Cryptography* 82, pp. 319–349.

Hermelin, Miia, Joo Yeon Cho, and Kaisa Nyberg (2008). «Multidimensional Linear Cryptanalysis of Reduced Round Serpent». In: *Information Security and Privacy, 13th Australasian Conference, ACISP 2008, Wollongong, Australia, July 7–9, 2008, Proceedings*. Ed. by Yi Mu, Willy Susilo, and Jennifer Seberry. Vol. 5107. LNCS. Springer, Berlin, Heidelberg, pp. 203–215. doi: 10.1007/978-3-540-70500-0_15. url: https://doi.org/10.1007/978-3-540-70500-0%5C_15.

Kaliski Jr., Burton S. and Matthew J. B. Robshaw (Aug. 1994). «Linear Cryptanalysis Using Multiple Approximations». In: CRYPTO'94. Ed. by Yvo Desmedt. Vol. 839. LNCS. Springer, Berlin, Heidelberg, pp. 26–39. doi: 10.1007/3-540-48658-54.

Vaudenay, Serge (1996a). «An Experiment on DES Statistical Cryptanalysis». In: *CCS '96, Proceedings of the 3rd ACM Conference on Computer and Communications Security, New Delhi, India, March 14–16, 1996*. Ed. by Li Gong and Jacques Stearn. ACM, New York, pp. 139–147. doi: 10.1145/238168.238206. url: https://doi.org/10.1145/238168.238206.

6.6. Упражнения

Упражнение 6.1

Пусть $\mu_1, \ldots, \mu_l$ – вещественные числа, не все равные нулю. Найти веса $w_1, \ldots, w_l$, для которых $\sum_{i=1}^{l} w_i^2 = 1$, такие что сумма $\sum_{i=1}^{l} w_i \mu_i$ максимальна. Каково ее максимальное значение?

Упражнение 6.2

Пусть $\mathbf{x}$ – случайная величина со средним μ и дисперсией σ^2.

1. Докажите, что если $\mathbf{x}$ – симметричная случайная величина, т. е. $-\mathbf{x}$ и $\mathbf{x}$ имеют одинаковое распределение, то

$$\mathbb{E}(\mathbf{x}^4) = \mathbb{E}((\mathbf{x} - \mu)^4 + 6\mu^2\sigma^2 + \mu^4.$$

2. Воспользовавшись этим результатом, завершите доказательство леммы 6.2.

Упражнение 6.3

Найдите обратное отношение в теореме 6.3.

Упражнение 6.4

Докажите следствие 6.5: квадратичное евклидово расхождение многомерной линейной аппроксимации Λ равно ее емкости.

Упражнение 6.5

Покажите, что если вероятности $p_1, \ldots, p_{|\Lambda|}$ известны, то информационная сложность критерия из раздела 6.2.2 пропорциональна величине, обратной квадратичному евклидову расхождению.

* Упражнение 6.6

Проанализируйте информационную сложность различителей, основанных на критерии Пирсона χ^2, в терминах квадратичного евклидова расхождения. Там, где уместно, опирайтесь на аппроксимации.

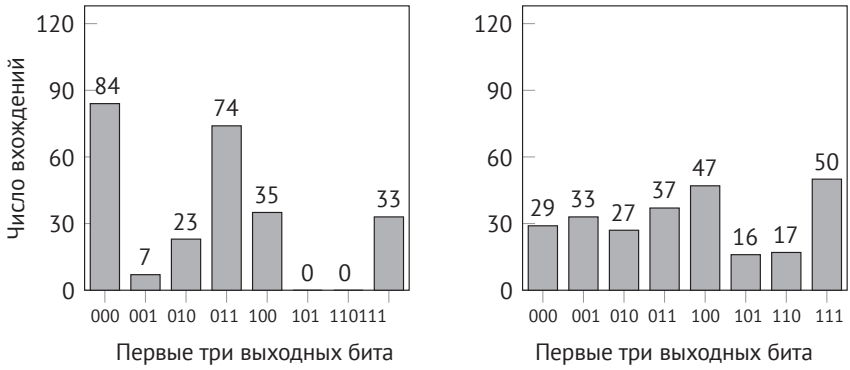
В терминах линейного криптоанализа объясните, почему статистика критерия χ^2 имеет $|\Lambda| - 1$ степеней свободы для $|\Lambda|$ категорий.

Упражнение 6.7

Блочный шифр $\mathsf{E}_k : \mathbb{F}_2^{128} \to \mathbb{F}_2^{128}$ был проанализирован с использованием линейного криптоанализа, что привело к следующим оценкам корреляций трех доминирующих линейных аппроксимаций:

$$C^{\mathsf{E}_k}_{01100\cdots0,\,10\cdots00} \approx (-1)^{k_1}\frac{1}{2} + (-1)^{k_2}\frac{1}{4},$$
$$C^{\mathsf{E}_k}_{10000\cdots0,\,10\cdots00} \approx (-1)^{k_1}\frac{1}{4} + (-1)^{k_3}\frac{1}{4},$$
$$C^{\mathsf{E}_k}_{11100\cdots0,\,10\cdots00} \approx (-1)^{k_1+k_3}\frac{1}{4}.$$

Чтобы убедиться в правильности этого анализа, были поставлены эксперименты для двух различных значений ключа. В каждом эксперименте по выборке (с возвращением) 256 открытых текстов вида «0 ∗ ∗ … ∗», где позиции, обозначенные *, выбираются независимо и равномерно из множества {0, 1}, вычисляется гистограмма значений первых трех выходных битов. Результаты показаны на рис. 6.3.



(a) Гистограмма для первого ключа (b) Гистограмма для второго ключа

Рис. 6.3. Число вхождений каждой комбинации первых трех битов шифртекста для выборки

1. Объясните, почему на рис. 6.3a и 6.3b число вхождений 100 и 111 почти одинаково.
2. Каково наиболее вероятное значение битов ключа k_1, k_2 и k_3 в эксперименте на рис. 6.3a?
3. Найдите значения k_1, k_2 и k_3 такие, что первые три бита шифртекста (почти) никогда не равны 000.

4. Нарисуйте эскиз наиболее вероятной гистограммы для эксперимента, основанного на том же ключе, что на рис. 6.3b, но для выборки (с возвращением) 256 открытых текстов вида «1 ∗ ∗ … ∗».

* Упражнение 6.8

Следующие вопросы ведут к общей атаке на построение Лая–Месси, показанное на рис. 6.4 для $n = 128$ бит. Ответы на первые четыре вопроса опровергают заявления о безопасности настраиваемого блочного шифра SPC[1]. Последний вопрос допускает различные ответы.

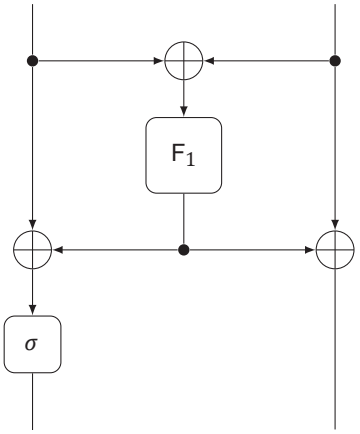


Рис. 6.4. Один раунд построения Лая–Месси. Функция $\sigma : \mathbb{F}_2^{64} \to \mathbb{F}_2^{64}$ определена как $\sigma(x_1 \| x_2) = x_2 \| (x_1 + x_2)$, где $x_1, x_2 \in \mathbb{F}_2^{32}$

Предположим, что случайные раундовые функции F_1, F_2, … независимы и равномерно распределены. Случайные функции более подробно исследуются в главе 7. В этом упражнении можно предполагать, что всякая нетривиальная линейная аппроксимация m-битовой случайной функции имеет корреляцию $\pm 2^{-m/2}$ со случайным равномерно распределенным знаком.

1. Найдите линейный след для трех раундов n-битового построения Лая–Месси с корреляцией, приближенно равной $\pm 2^{-n/4}$.
2. Опишите трехраундовый многомерный линейный различитель, использующий приблизительно $2^{n/4}$ данных. Сопоставьте его с χ^2-различителем.
3. Используйте выбранные открытые тексты, чтобы распространить свой различитель на четыре раунда с использованием того же объема данных. В этом различителе кое-что необычно, но почему он все равно будет работать для такого шифра, как SPC?
4. Придумайте атаку с частичным восстановлением сообщения. То есть получите *какую-нибудь* новую информацию об открытых текстах. Можете предполагать, что открытые тексты уже частично известны противнику.
5. Обозначим первую раундовую функцию Лая–Месси F_1. Предложите метод получения выхода F_1 для выбранных входов с использованием при-

[1] *https://github.com/veorq/spc*.

мерно такого же объема данных, как для атаки с частичным восстановлением сообщения.

6. В шифре SPC функция F_1 определена на основе криптографической функции `SipHash-1-2`. Предложите атаку с восстановлением ключа с выбранной настройкой и с выбранным открытым текстом для F_1. Выведите атаку с восстановлением ключа для шифра SPC с полным числом раундов.

   Указание: для ответа на этот вопрос одного линейного криптоанализа может не хватить.

Глава 7

Оптимальная проверка статистических гипотез

В предыдущих главах и в главах 4 и 6 в особенности мы рассмотрели методы проверки статистических гипотез. Мы использовали статистические критерии, чтобы определить, соответствует ли заданная эмпирическая корреляция реальному или неправильному ключу. В этой главе мы более систематически подойдем к проверке статистических гипотез и выведем методы, которые – в некотором весьма определенном смысле – являются наилучшими из возможных.

7.1. Вероятностные меры

Большинство результатов этой главы применимы как к дискретным, так и к непрерывным распределениям вероятностей. Чтобы избежать повторов, удобно воспользоваться языком теории меры. Для чтения этой главы не требуется предварительного знакомства с материалом, при условии что читатель будет иметь в виду следующие замечания.

Пространством с мерой называется тройка $(X, \mathfrak{S}, \mu)$, где X – непустое множество, $\mathfrak{S}$ – множество подмножеств X, содержащее X, замкнутое относительно дополнения и счетного объединения, а μ – мера. Мерой называется функция $\mu: \mathfrak{S} \to \mathbb{R}$, которая принимает неотрицательные значения, удовлетворяет условию $\mu(\emptyset) = 0$ и является *счетно-аддитивной*: $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ для непересекающихся множеств $A_1, A_2, \ldots \in \mathfrak{S}$. Элементы $\mathfrak{S}$ называются измеримыми множествами.

Вещественная функция $f : X \to \mathbb{R}$ называется измеримой, если прообразы всех интервалов вида $(-\infty, t)$ измеримы. В этом случае можно определить интеграл f по измеримому множеству Y, обозначаемый

$$\int_Y f(x)\,\mu(dx).$$

Технические детали определения таких интегралов намеренно оставлены за кадром. Для целей данной главы достаточно понимать два примера, приведенных в конце этого раздела. Иногда мы будем говорить, что две измеримые функции равны почти всюду (п. в.), если они различаются на множестве меры нуль.

Мера $P : \mathfrak{S} \to \mathbb{R}$, удовлетворяющая условию $P(X) = 1$, называется *вероятностной*. С точки зрения теории меры, распределения вероятностей являются вероятностными мерами. Если существует измеримая функция $p: X \to \mathbb{R}$ такая, что для любого измеримого множества Y

$$P(Y) = \int_Y p(x)\,\mu(dx),$$

то P называется *абсолютно непрерывной* относительно μ. Функция p называется *плотностью Радона–Никодима* функции P.

Пример 7.1. Если X – конечное множество и $\mathfrak{S}$ – множество всех подмножеств X, то μ можно выбрать так, что она будет *считающей мерой*: $\mu(Y) = |Y|$ для всех $Y \in \mathfrak{S}$. Любое распределение вероятностей $P : \mathfrak{S} \to \mathbb{R}$ абсолютно непрерывно относительно считающей меры и

$$P(Y) = \int_Y p(x)\,\mu(dx) = \sum_{x \in Y} p(x).$$

Плотность Рандона–Никодима $p(x) = P(\{x\})$ – это хорошо знакомая функция массы вероятности. ▷

Пример 7.2. Для $X = \mathbb{R}$ построение $\mathfrak{S}$ технически несколько сложнее – достаточно сказать, что оно содержит все вещественные интервалы $Y = (a, b)$. Стандартная мера на $\mathbb{R}$ называется мерой Лебега, для нее имеет место равенство $\mu(Y) = b - a$. Если P – вероятностная мера, то ее функция распределения имеет вид $x \mapsto P((-\infty, x])$. Если функция распределения дифференцируема, то P имеет плотность:

$$P(Y) = \int_Y p(x)\,\mu(dx) = \int_a^b p(x)\,dx.$$

Интеграл в правой части – это стандартный интеграл Лебега на $\mathbb{R}$. ▷

7.2. Простые гипотезы

В данном разделе мы рассмотрим проверки так называемых *простых* гипотез. В этом случае имеется наблюдение x и две гипотезы ① и ② о распределении вероятностей, из которого оно было выбрано:

Гипотеза ①: наблюдение x было выбрано из распределения P_1.

Гипотеза ②: наблюдение x было выбрано из распределения P_2.

Тонкое, но важное отличие от того, что мы делали в главе 4, заключается в том, что гипотезы ① и ② должны полностью определять распределения P_1 и P_2. Например, гипотезы, определяющие только средние P_1 и P_2, не являются *простыми*.

В контексте линейных атак с восстановлением ключа наблюдение x является либо вектором эмпирических корреляций, либо эмпирическим распределением линейной проекции наблюдаемых пар (открытый текст, шифртекст) (мно-

гомерный линейный криптоанализ). Гипотеза ① соответствует правильному ключу, а гипотеза ② – неправильному. Следовательно, эти гипотезы являются простыми, только когда выполнены два важных предположения: (i) все корреляции известны и (ii) для неправильных ключей образцы случайны и равномерно распределены. Второе предположение является частью «простой модели» из главы 4. В разделе 7.3 обсуждается, что бывает, когда одно из этих предположений не выполняется.

7.2.1. Теория Неймана–Пирсона

Хотя в названии этой главы говорится об *оптимальной проверке*, мы еще не определили, что это значит. Напомним (см. главу 4), что у любой проверки гипотез имеется вероятность успеха P_S и вероятность ложноположительного результата P_F. Компромисс между P_S и P_F обсуждался Нейманом и Пирсоном, которые называли вероятности $1 - P_S$ и P_F частотами ошибок первого и второго рода. С обоими типами ошибок обычно ассоциируется некоторая стоимость. В общем случае она описывается функцией стоимости $f(1 - P_S, P_F)$, возрастающей по обеим переменным. Например, в случае линейной атаки с восстановлением ключа функцией стоимости может быть временная или информационная сложность атаки.

На первый взгляд, из вышесказанного может сложиться впечатление, что не существует одного критерия, который минимизировал бы любую функцию стоимости $f(1 - P_S, P_F)$. Однако Нейман и Пирсон показали, что такой критерий есть. Точнее, существует критерий, который минизирует вероятность ложноположительного результата для любого выбора вероятности успеха. Такой критерий называется *равномерно наиболее мощным*, потому что $1 - P_F$ называют также мощностью критерия.

Для любой проверки простых гипотез существует измеримое множество $\mathcal{A}$ («область принятия гипотезы») такое, что гипотеза ① принимается, когда $x \in \mathcal{A}$, а гипотеза ② – в противном случае. Отсюда следует, что $P_S = P_1(\mathcal{A})$ и $P_F = P_2(\mathcal{A})$. Лемма Неймана–Пирсона дает множества $\mathcal{A}$ такие, что P_F минимально для данного P_S.

Теорема 7.1 *(лемма Неймана–Пирсона). Пусть P_1 и P_2 – вероятностные меры на пространстве с мерой $(X, \mathfrak{S}, \mu)$ с плотностями p_1 и p_2 (в смысле Радона–Никодима). Для любого вещественного $\tau > 0$ обозначим*

$$\mathcal{A}_\tau = \Big\{ x \in X \mid p_1(x) > \tau p_2(x) \Big\}.$$

Если $\mathcal{B}$ – измеримое множество такое, что $P_1(\mathcal{B}) \geqslant P_1(\mathcal{A}_\tau)$, то $P_2(\mathcal{B}) \geqslant P_2(\mathcal{A}_\tau)$.

Доказательство. Из определения $\mathcal{A}_\tau$ вытекают следующие неравенства:

$$P_2(\mathcal{B} \setminus \mathcal{A}_\tau) = \int_{\mathcal{B}\setminus\mathcal{A}_\tau} p_2(x)\,\mu(dx) \geq \frac{1}{\tau} \int_{\mathcal{B}\setminus\mathcal{A}_\tau} p_1(x)\,\mu(dx) = \frac{1}{\tau} P_1(\mathcal{B}\setminus\mathcal{A}_\tau);$$

$$P_1(\mathcal{A}_\tau \setminus \mathcal{B}) = \int_{\mathcal{A}_\tau\setminus\mathcal{B}} p_1(x)\,\mu(dx) \geq \tau \int_{\mathcal{A}_\tau\setminus\mathcal{B}} p_2(x)\,\mu(dx) = \tau\ P_2(\mathcal{A}_\tau\setminus\mathcal{B}).$$

Отсюда

$$P_2(\mathcal{B}) = P_2(\mathcal{A}_\tau \cap \mathcal{B}) + P_2(\mathcal{B} \setminus \mathcal{A}_\tau) \geq P_2(\mathcal{A}_\tau \cap \mathcal{B}) + \frac{1}{\tau} P_1(\mathcal{B} \setminus \mathcal{A}_\tau).$$

Из условия $P_1(\mathcal{B}) \geqslant P_1(\mathcal{A}_\tau)$ следует, что $P_1(\mathcal{B} \setminus \mathcal{A}_\tau) \geqslant P_1(\mathcal{A}_\tau \setminus \mathcal{B})$. Подстановка этого неравенства в правую часть дает

$$P_2(\mathcal{B}) \geq P_2(\mathcal{A}_\tau \cap \mathcal{B}) + \frac{1}{\tau} P_1(\mathcal{A}_\tau \setminus \mathcal{B}) \geq P_2(\mathcal{A}_\tau \cap \mathcal{B}) + P_2(\mathcal{A}_\tau \setminus \mathcal{B}) = P_2(\mathcal{A}_\tau).$$

Следовательно, $P_2(\mathcal{B}) \geqslant P_2(\mathcal{A}_\tau)$, что и требовалось доказать. □

Область принятия гипотезы $\mathcal{A}_\tau$, определенная в теореме 7.1, соответствует критерию, который сравнивает статистику критерия отношения правдоподобия t_{lr} с пороговой величиной τ. Эта статистика критерия определяется как

$$t_{\text{lr}}(x) = \frac{p_1(x)}{p_2(x)}.$$

На практике обычно используют логарифм t_{lr} – это эквивалентно, потому что логарифм является возрастающей функцией. Результирующая статистика критерия называется *логарифмическим отношением правдоподобия* t_{llr}:

$$t_{\text{llr}}(x) = \log \frac{p_1(x)}{p_2(x)}.$$

Хотя теорема 7.1 показывает, что критерий (логарифмического) отношения правдоподобия является равномерно наиболее мощным, она не дает значений P_{S} и P_{F}. В следующих двух разделах эти значения определяются в двух частных случаях: когда P_1 и P_2 – многомерные нормальные распределения и когда P_1 и P_2 почти равны.

7.2.2. Два многомерных нормальных распределения

В множественном линейном криптоанализе эмпирические корреляции приблизительно нормально распределены, когда число образцов q достаточно велико. В этом разделе исследуется критерий отношения правдоподобия для случая многомерных нормальных P_1 и P_2.

Предположим, что P_1 и P_2 – многомерные нормальные распределения со средними μ_1 и μ_2 соответственно и одинаковой ковариационной матрицей Σ размера $l \times l$. В случае множественного линейного криптоанализа μ_1 – вектор известных корреляций, а $\mu_2 = 0$ (простая модель). В теореме 6.1 было показано, что $\Sigma \approx I/q$ часто является хорошим приближением в данном случае. Для многомерных линейных атак, основанных на эмпирическом распределении вероятностей линейной проекции образцов, μ_1 содержит истинные вероятности, а $\mu_2 \equiv 1/l$. Заметим, что P_1 и P_2 вырождены в многомерном случае из-за включения тривиальной аппроксимации (0, 0), или, эквивалентно, потому что сумма эмпирических вероятностей равна 1. Опустив тривиальную аппроксимацию или одну из эмпирических вероятностей, мы решим эту проблему.

Плотности распределений вероятностей P_1 и P_2 удовлетворяет соотношению

$$p_i(x) \propto \exp\Big(-\frac{1}{2}(x-\mu_i)^\mathsf{T}\Sigma^{-1}(x-\mu_i)\Big).$$

Следовательно, с точностью до постоянного множителя логарифмическое отношение правдоподобия равно

$$\begin{aligned} t_{\mathsf{llr}} &= (x-\mu_2)^\mathsf{T}\Sigma^{-1}(x-\mu_2) - (x-\mu_1)^\mathsf{T}\Sigma^{-1}(x-\mu_1) \\ &= 2(\mu_1-\mu_2)^\mathsf{T}\Sigma^{-1}x + \mu_2^\mathsf{T}\Sigma^{-1}\mu_2 - \mu_1^\mathsf{T}\Sigma^{-1}\mu_1. \end{aligned}$$

С точностью до сдвига и масштабирования статистика критерия $\mathbf{t}_{\mathsf{lda}} = (\mu_1 - \mu_2)^\mathsf{T}\Sigma^{-1}\mathbf{x}$. Это можно переписать в виде линейной комбинации элементов $\mathbf{x}$:

$$\mathbf{t}_{\mathsf{lda}} = \sum_{i=1}^{l} w_i\,\mathbf{x}_i.$$

Среднее этой статистики критерия равно $(\mu_1 - \mu_2)^\mathsf{T}\Sigma^{-1}\mu_1$ в случае истинности гипотезы ① и $(\mu_1 - \mu_2)^\mathsf{T}\Sigma^{-1}\mu_2$ в случае истинности гипотезы ②. Дисперсия равна $(\mu_1 - \mu_2)^\mathsf{T}\Sigma^{-1}(\mu_1 - \mu_2)$ – см. приложение А. Выбор $w \propto (\mu_1 - \mu_2)^\mathsf{T}\Sigma^{-1}$ максимизирует разность между средними для гипотез ① и ②, сохраняя дисперсию постоянной. Это в точности тот подход, который мы использовали в разделе 6.1.2.

Метод, обсуждавшийся до сих пор, в литературе по статистике называется *линейным дискриминантным анализом*. У него есть простая геометрическая интерпретация: распределения P_1 и P_2 разделены гиперплоскостью, ортогональной вектору $(\mu_1 - \mu_2)^\mathsf{T}\Sigma^{-1}$. Это показано на рис. 7.1 для $\Sigma \propto I$. Если ковариационные матрицы P_1 и P_2 не равны, то оптимальным критерием является *квадратичный дискриминантный анализ*.

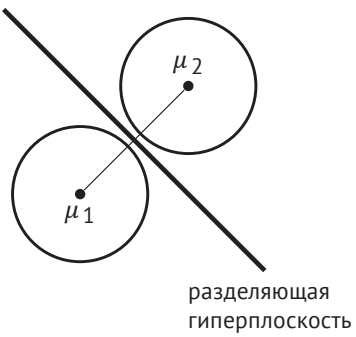


Рис. 7.1. Разделение P_1 и P_2 прямой линией

7.2.3. Два распределения почти равны

Анализ в разделе 7.2.2 показывает, что критерии с известной корреляцией из разделов 6.1.2 и 6.2.2 являются равномерно наиболее мощными в простой модели с дополнительными аппроксимациями из раздела 6.1.2. Однако это предполагает, что различитель основан на векторе эмпирических корреляций (или эмпирических вероятностей в многомерном случае). В многомерном случае

можно сделать еще один шаг и дать различителю прямой доступ к линейным проекциям $\pi_\Lambda \perp (z_i)$ образцов $z_1, \ldots, z_q$.

Если образцы независимы и одинаково распределены, то P_1 и P_2 будут q-кратными произведениями распределений R_1 и R_2. В частности, если r_1 и r_2 – плотности R_1 и R_2 соответственно, то $p_i(x_1, \ldots, x_q) = \prod_{j=1}^{q} r_i(x_j)$. Отсюда логарифмическое отношение правдоподобия равно

$$\mathbf{t}_{\text{llr}} = \log \frac{p_1(\mathbf{x}_1, \ldots, \mathbf{x}_q)}{p_2(\mathbf{x}_1, \ldots, \mathbf{x}_q)} = \sum_{i=1}^{q} \log \frac{r_1(\mathbf{x}_i)}{r_2(\mathbf{x}_i)}.$$

Поскольку наблюдения $\mathbf{x}_1, \ldots, \mathbf{x}_q$ независимы, центральная предельная теорема говорит, что распределение $\mathbf{t}_{\text{llr}}/\sqrt{q}$ сходится к нормальному. Следовательно, асимптотическая информационная сложность критерия отношения правдоподобия определяется теоремой 4.1.

Если истинна гипотеза ①, то среднее $\mathbf{t}_{\text{llr}}/q$ равно

$$I_{1:2} = \int r_1(x) \log \frac{r_1(x)}{r_2(x)}\, \mu(dx).$$

Аналогично, если истинна гипотеза ②, то среднее равно $-I_{2:1}$, где

$$I_{2:1} = \int r_2(x) \log \frac{r_2(x)}{r_1(x)}\, \mu(dx).$$

Кульбак и Лейблер[1] называли $I_{1:2}$ и $I_{2:1}$ *средней информацией дискриминации* между гипотезами ①-② или ②-① соответственно. Они также определили $J_{12} = I_{1:2} + I_{2:1}$ как расхождение между R_1 и R_2. В наши дни $I_{1:2}$ называется *расхождением Кульбака–Лейблера* R_1 относительно R_2, а J_{12} – *расхождением Джеффриса* между R_1 и R_2.

Поскольку J_{12} – разность между средними статистики критерия $\mathbf{t}_{\text{llr}}/q$ при условии истинности гипотез ① и ②, она играет важную роль для определения вероятностей успеха и ложноположительного результата в критерии отношения правдоподобия. Однако точное соотношение зависит также от дисперсии $\mathbf{t}_{\text{llr}}/q$. Если истинна гипотеза ①, то дисперсия равна $(V_{1:2} - I_{1:2}^2)/q$, где

$$V_{1:2} = \int r_1(x) \left(\log \frac{r_1(x)}{r_2(x)} \right)^2 \mu(dx).$$

Аналогично, если истинна гипотеза ②, то дисперсия равна $(V_{2:1} - I_{2:1}^2)/q$, где

$$V_{2:1} = \int r_2(x) \left(\log \frac{r_2(x)}{r_1(x)} \right)^2 \mu(dx).$$

Если распределения R_1 и R_2 близки, то вычисления упрощает следующая лемма.

Лемма 7.2. *Пусть r_1 и r_2 – почти всюду ненулевые плотности вероятности относительно общей вероятностной меры μ, и определим $I_{1:2}$, $I_{2:1}$, $V_{1:2}$ и $V_{2:1}$, как показано выше. Если $|r_1(x) - r_2(x)| \leqslant \epsilon \min\{|r_1(x)|, |r_2(x)|\}$ почти всюду, то $I_{2:1} = I_{1:2} + O(\epsilon^3)$ и*

[1] Соломон Кульбак и Ричард Лейблер работали криптоаналитиками в АНБ.

$$I_{1:2} = \frac{1}{2}\int \frac{\big(r_1(x) - r_2(x)\big)^2}{r_2(x)}\mu(dx) + \mathcal{O}(\epsilon^3).$$

Кроме того, $V_{2:1} = V_{1:2} + O(\epsilon^3)$ *и* $V_{1:2} = 2\, I_{1:2} + O(\epsilon^3)$.

Доказательство. Если $\epsilon_{1:2}(x) = (r_2(x) - r_1(x))/r_1(x)$, то

$$\begin{aligned} I_{1:2} &= -\int r_1(x)\log\big(1+\varepsilon_{1:2}(x)\big)\mu(dx) \\ &= \frac{1}{2}\int r_1(x)\,\varepsilon_{1:2}^2(x)\,\mu(dx) - \underbrace{\int \big(r_1(x) - r_2(x)\big)\mu(dx)}_{0} + \mathcal{O}(\epsilon^3). \end{aligned}$$

Второе из равенств выше следует из разложения в ряд Тейлора $t \mapsto \log(1 + t)$ в точке $t = 0$. Второй член обращается в нуль, потому что интегралы r_1 и r_2 равны 1. Аналогично, полагая $\epsilon_{2:1}(x) = (r_1(x) - r_2(x))/r_2(x)$, получаем

$$I_{2:1} = \frac{1}{2}\int r_2(x)\,\varepsilon_{2:1}^2(x)\,\mu(dx) + \mathcal{O}(\epsilon^3).$$

Желаемый результат следует из того, что

$$\frac{\big(r_1(x) - r_2(x)\big)^2}{r_2(x)} = \frac{\big(r_1(x) - r_2(x)\big)^2}{r_1(x)}\underbrace{\frac{1}{1+\varepsilon_{1:2}(x)}}_{1+\mathcal{O}(\epsilon)} = \frac{\big(r_1(x) - r_2(x)\big)^2}{r_1(x)} + \mathcal{O}(\epsilon^3)r_1(x).$$

Что касается второго утверждения, мы сначала покажем, что $V_{1:2} - V_{2:1} = O(\epsilon^3)$:

$$V_{1:2} - V_{2:1} = \int r_2(x)\underbrace{\varepsilon_{2:1}(x)\log^2\big(1+\varepsilon_{2:1}(x)\big)}_{\mathcal{O}(\epsilon^3)}\mu(dx) = \mathcal{O}(\epsilon^3).$$

Теперь достаточно показать, что $V_{2:1} = 2I_{2:1} + O(\epsilon^3)$. В силу разложения $t \mapsto \log(1 + t)$ в ряд Тейлора в точке $t = 0$, имеем

$$\left(\log\frac{r_2(x)}{r_1(x)}\right)^2 = 2\log\frac{r_2(x)}{r_1(x)} + 2\varepsilon_{2:1}(x) + \mathcal{O}(\epsilon^3).$$

Следовательно, $V_{2:1}$ удовлетворяет соотношению

$$V_{2:1} = \int r_2(x)\left(\log\frac{r_2(x)}{r_1(x)}\right)^2\mu(dx) = 2I_{2:1} + 2\underbrace{\int \big(r_1(x) - r_2(x)\big)\mu(dx)}_{0} + \mathcal{O}(\epsilon^3).$$

Второй член обращается в нуль, потому что интегралы r_1 и r_2 равны 1. □

Лемма 7.2 показывает, что дисперсия $\mathbf{t}_{\text{llr}}/q$ одинакова как для гипотезы ①, так и для гипотезы ② с точностью до погрешности $O(\epsilon_3)$. Точнее, дисперсия приближенно равна разности между средними, поделенной на q. На рис. 7.2 показана ситуация в целом. По теореме 4.1, количество образцов q равно

$$q = \frac{\left(\Phi^{-1}(P_S) - \Phi^{-1}(P_F)\right)^2}{2\, I_{1:2}}$$

в предположении, что $P_S \geqslant P_F$. Следовательно, информационная сложность обратно пропорциональна расхождению Кульбака–Лейблера $I_{1:2}$. Для многомерной линейной аппроксимации $\Lambda \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^m$ распределение R_1 дискретно на $(\mathbb{F}_2^n \times \mathbb{F}_2^m)/\Lambda^\perp$, а R_2 равномерно на том же множестве. Для этого случая лемма 7.2 дает следующую аппроксимацию $2\, I_{1:2}$:

$$2\, I_{1:2} = |\Lambda| \sum_z \left(r_1(z) - \frac{1}{|\Lambda|} \right)^2,$$

где суммирование производится по всем $z \in (\mathbb{F}_2^n \times \mathbb{F}_2^m)/\Lambda^\perp$. Это в точности квадратичное евклидово расхождение, которое, в силу следствия 6.5, равно $\mathrm{Cap}(\Lambda)$.

7.3. Составные гипотезы

Чаще всего корреляции линейных аппроксимаций зависят от ключа. Аналогично для неправильно угаданного ключа средние эмпирические корреляции являются (как правило) малыми зависящими от ключа значениями, но не нулем. Это означает, что, в противоположность разделу 7.2, гипотезы ① и ② не полностью определяют распределения P_1 и P_2 – они не являются простыми гипотезами.

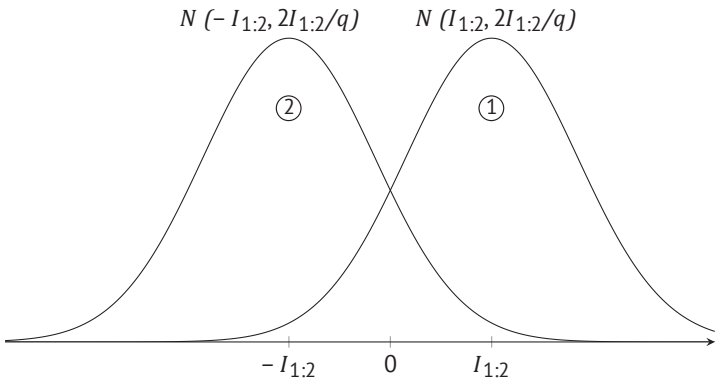


Рис. 7.2. Асимптотическое распределение статистики критерия логарифмического отношения правдоподобия при гипотезах ① и ②

В общем случае если задано наблюдение x, то в задаче о проверке составной гипотезы требуется решить, из какого из двух семейств распределений была произведена выборка x. То есть налицо следующие две *составные* гипотезы:

Гипотеза ①: x было выбрано из распределения, принадлежащего семейству $\{P_1^{\theta_1} | \ \theta_1 \in \Theta_1\}$.

Гипотеза ②: x было выбрано из распределения, принадлежащего семейству $\{P_2^{\theta_2} | \ \theta_2 \in \Theta_2\}$.

$P_1^{\theta_1}$ и $P_2^{\theta_2}$ можно представлять себе как параметризованные распределения. Чтобы задача имела смысл, семейства $\{P_1^{\theta_1} \mid \theta_1 \in \Theta_1\}$ и $\{P_2^{\theta_2} \mid \theta_2 \in \Theta_2\}$ не обязаны быть непересекающимися при условии, что для множеств Θ_1 и Θ_2 известны априорные распределения вероятностей. Например, если Θ_1 состоит из всех возможных значений битов ключа, от которых зависит корреляция, то априорным является равномерное распределение на Θ_1. Априорные распределения на Θ_1 и Θ_2 будем называть *гипотезой рандомизации с правильным ключом* и *гипотезой рандомизации с неправильным ключом* соответственно. Их выбор обсуждается в разделах 7.3.2 и 7.3.3.

Результаты из раздела 7.2 неприменимы к составным гипотезам. В частности, необязательно существует равномерно наиболее мощный критерий. Тем не менее можно найти критерии, которые минимизируют среднее $\mathbb{E}f(1 - P_S(\boldsymbol{\theta}_1), P_F(\boldsymbol{\theta}_2))$ конкретной функции стоимости f, где среднее берется относительно априорных распределений на Θ_1 и Θ_2. Эта проблема обсуждается в разделе 7.3.1.

7.3.1. Коэффициенты Байеса

В отсутствие равномерно наиболее мощного критерия мы можем рассмотреть критерии, наиболее мощные в среднем. То есть для всех средних вероятностей успеха $\mathbb{E}(P_S(\boldsymbol{\theta}_1))$ такой критерий должен минимизировать среднюю вероятность ложноположительного результата $\mathbb{E}(P_S(\boldsymbol{\theta}_2))$. Это эквивалентно критерию различения простых гипотез, соответствующих апостериорным распределениям P_1 и P_2, плотности которых p_1 и p_2 определяются усреднением по параметрам:

$$p_i(x) = \int_{\Theta_i} p_i^{\theta}(x)\, q_i(\theta)\, \mu(d\theta),$$

где p_i^{θ} – плотность p_i^{θ}, а q_i – плотность априорного распределения.

Используя результаты из раздела 7.2, статистика критерия отношения правдоподобия дает равномерно наиболее мощный критерий (в среднем, как было сказано выше):

$$t_{\mathrm{lr}}(x) = \frac{p_1(x)}{p_2(x)} = \frac{\int p_1^{\theta}(x)\, q_1(\theta)\, \mu(d\theta)}{\int p_2^{\theta}(x)\, q_2(\theta)\, \mu(d\theta)}.$$

Эту величину называют также коэффициентом Байеса для гипотез ① и ②.

7.3.2. Гипотеза рандомизации с правильным ключом

Априорное распределение на Θ_1 вытекает непосредственно из анализа шифра, который приводит к зависящей от ключа аппроксимации каждой корреляции. Например, пусть Λ – множественная линейная аппроксимация. Анализ приводит к множеству классов ключей $\mathcal{K}$ такому, что корреляции для класса ключей $k \in \mathcal{K}$ определяются известным вектором $\mu_k \in \mathbb{R}^{|\Lambda|}$.

Если предположить, что ключи имеют равномерное априорное распределение, то априорную вероятность f_k каждого класса ключей $k \in \mathcal{K}$ можно вычислить или, если развертка ключа сложна, оценить. Внутри каждого класса ключей распределение эмпирических корреляций является многомерным нормальным со средним μ_k и ковариационной матрицей I/q (приближенно, как показывает теорема 6.1). Следовательно, плотность вероятности p_1 пропорциональна

$$p_1(x) \propto \sum_{k\in\mathcal{K}} f_k \exp\Big(-\frac{q}{2}(x-\mu_k)^\intercal(x-\mu_k)\Big).$$

Такое распределение называется смесью многомерных нормальных распределений. В простой модели эмпирические корреляции имеют нулевое среднее и ковариационную матрицу I/q при условии истинности гипотезы ②. Следовательно, отношение правдоподобия пропорционально

$$\frac{p_1(x)}{p_2(x)} \propto \sum_{k\in\mathcal{K}} f_k \exp\Big(q\,\mu_k^\intercal x - \frac{q}{2}\,\mu_k^\intercal\mu_k\Big).$$

В общем случае не существует «элементарного» замкнутого выражения информационной сложности этого критерия. Даже если емкость (почти) не зависит от ключа, информационная сложность может быть пропорциональна $1/\mathrm{Cap}(\Lambda)$, $\sqrt{|\Lambda|}/\mathrm{Cap}(\Lambda)$ или чему-то промежуточному.

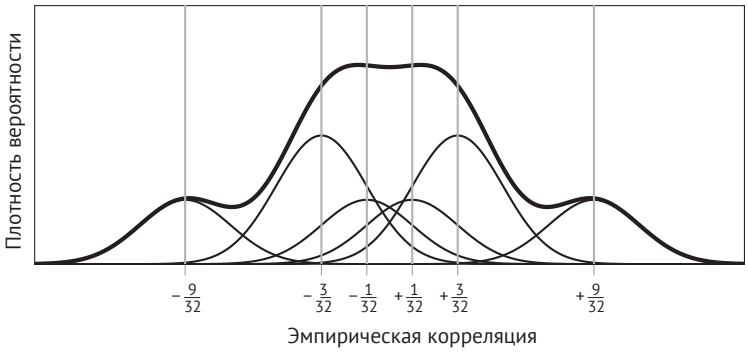


Рис. 7.3. Зависимость плотности вероятности от эмпирической корреляции

Пример 7.3. Для линейной аппроксимации из примера 2.3 с корреляцией $(-1)^{\kappa_1}/8\ (1 + (-1)^{\kappa_2}/2)(1 + (-1)^{\kappa_3}/2)$ апостериорное распределение P_1 является смесью шести нормальных распределений со средними $\pm1/32$, $\pm3/32$ и $\pm9/32$. Функции плотности вероятности p_1 распределения P_1 и его компонент показаны на рис. 7.3 для $q = 256$ образцов. С точностью до постоянного множителя отношение правдоподобия равно

$$e^{-\frac{1}{2}q(9/32)^2}\cosh\left(\frac{9qx}{32}\right)+2e^{-\frac{1}{2}q(3/32)^2}\cosh\left(\frac{3qx}{32}\right)+e^{-\frac{1}{2}q(1/32)^2}\cosh\left(\frac{qx}{32}\right).$$

Для больших q в отношении правдоподобия преобладает член, соответствующий корреляции $\pm1/32$. Это объясняется экспоненциальными множителями в каждом члене, наибольший из которых равен $\exp(-q/32^2)$. Иными словами, поскольку от атаки с высокой средней вероятностью успеха требуется, чтобы она работала для большинства ключей, вероятность ложноположительного результата для нее в основном определяется ключами с низкой корреляцией. ▷

В частных случаях можно получить замкнутые (приближенные) формулы для информационной сложности. Один из таких случаев – когда от ключа за-

висят только знаки корреляций. Например, положим $l = |\Lambda|$ и предположим, что существуют такие положительные постоянные $c_1, \ldots, c_l$, что

$$\mu_k = \begin{bmatrix} (-1)^{k_1} c_1 \\ (-1)^{k_2} c_2 \\ \vdots \\ (-1)^{k_l} c_l \end{bmatrix}.$$

Кроме того, предположим, что $f_k = 1/2^l$ для любого k. Отношение правдоподобия пропорционально

$$\frac{p_1(x)}{p_2(x)} \propto \sum_{k \in \mathbb{F}_2^l} \prod_{i=1}^{l} e^{q(-1)^{k_i} c_i x_i} = \prod_{i=1}^{l} \frac{e^{q c_i x_i} + e^{-q c_i x_i}}{2} = \prod_{i=1}^{l} \cosh(q c_i x_i).$$

Если l велико, то можно ожидать, что qc_i^2 мало. Поскольку наблюдения x_i являются оценками корреляции, $qc_i x_i$ также мало. Следовательно, асимптотически при $qc_i x_i \to 0$ логарифмическое отношение правдоподобия равно (с точностью до постоянной C)

$$\log \frac{p_1(x)}{p_2(x)} + C = \sum_{i=1}^{l} \log \cosh(q c_i x_i) \sim \frac{1}{2} \sum_{i=1}^{l} q^2 c_i^2 x_i^2.$$

То есть логарифмическое отношение правдоподобия хорошо аппроксимируется взвешенной суммой квадратов оценок корреляций, в которой сами веса пропорциональны квадратам корреляций. Это в точности статистика критерия, которую мы использовали в разделе 6.1.2, с информационной сложностью, пропорциональной

$$\frac{1}{\sqrt{\sum_{i=1}^{l} c_i^4}}.$$

Следует помнить, что эта информационная сложность оптимальна, только когда априорное распределение ключа равномерно.

Наконец, заметим, что точные формулы корреляций редко бывают доступны. Поэтому ошибки модели в общем случае неизбежны. Это можно принять во внимание, модифицировав априорное распределение. Например, можно включить нормально распределенную ошибку с нулевым средним. Это полезно для противодействия чрезмерной уверенности модели и может рассматриваться как форма регуляризации.

7.3.3. Гипотеза рандомизации с неправильным ключом

Как обсуждалось в разделе 7.3.2, гипотеза ① часто является составной, потому что корреляции зависят от ключа. На практике корреляции для неправильно угаданных ключей тоже не равны в точности нулю. Поэтому гипотеза ② также должна быть составной.

В принципе, можно определить приближенные зависящие от ключа выражения для корреляций, соответствующих неправильно угаданным клю-

чам. Однако для этого необходим дополнительный анализ шифра, включая внешние раунды восстановления ключей, которые не принимались во внимание в простой модели. Статистический анализ концептуально такой же, как в разделе 7.3.2.

В отсутствие детального анализа существует также более общее составное уточнение гипотезы ②. Если частичное шифрование и дешифрирование достаточно сложны, то можно предположить, что они похожи на случайные перестановки, когда ключ неправильный. Это приводит к модели случайной перестановки, которая говорит, что априорное распределение корреляций при условии истинности гипотезы ② такое же, как для случайной перестановки.

Распределение корреляции линейной аппроксимации случайной функции или перестановки дает следующий результат. Сходимость в теореме 7.3 быстрая, если $|\Lambda|$ не слишком велико, поэтому обычно она дает хорошую аппроксимацию. Доказательство этого результата – предмет упражнений 7.1 и 7.2.

Теорема 7.3. *Пусть* $\mathbf{F}$ *– равномерно распределенная случайная функция или перестановка* $\mathbb{F}_2^n \to \mathbb{F}_2^m$. *Пусть* $\Lambda = \{(u_1, v_1), \ldots, (u_l, v_l)\} \subset \mathbb{F}_2^n \times \mathbb{F}_2^m$ *– множественная линейная аппроксимация* $\mathbf{F}$ *такая, что* $(0, 0) \notin \Lambda$. *Распределение вероятностей случайного вектора корреляций*

$$\sqrt{2^n} \begin{bmatrix} C^{\mathbf{F}}_{v_1, u_1} \\ C^{\mathbf{F}}_{v_2, u_2} \\ \vdots \\ C^{\mathbf{F}}_{v_l, u_l} \end{bmatrix}$$

сходится к многомерному нормальному распределению $\mathcal{N}(0, I)$ при $n \to \infty$.

Тогда для множественного линейного криптоанализа гипотеза ② заключается в том, что эмпирические корреляции имеют многомерное нормальное распределение $\mathcal{N}(\theta, I/q)$. Априорным распределением на Θ_2 является $\theta \sim \mathcal{N}(0, I/2^n)$. Зная функцию плотности вероятности, нетрудно видеть, что апостериорным распределением при этом будет $\mathcal{N}(0, I/q + I/2^n)$. Для одномерного случая это показано на рис. 7.4.

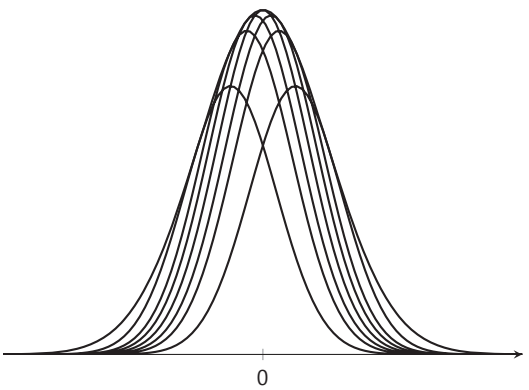


Рис. 7.4. Гипотеза с неверным ключом, основанная на модели случайной перестановки

Эффект модели случайной перестановки заключается в том, что она увеличивает дисперсию апостериорного распределения при условии истинности гипотезы ②. Это также означает, что дисперсии обоих апостериорных распределений различны. Для выборки без возвращения дисперсия при условии истинности гипотезы ② с известной корреляцией равна $1/q\,(1 - q/2^n) + 1/2^n = 1/q$.

Хотя уточненная гипотеза с неверным ключом, основанная на модели случайной перестановки, часто оказывает несущественное влияние на оценки стоимости атак, особенно когда $q \ll 2^n$, она ведет к важным выводам. В сочетании с разделом 7.3.2 она лучше объясняет трудность использования линейных аппроксимаций с абсолютной корреляцией c, меньше $2^{-n/2}$. Наивное объяснение состоит в том, что поскольку информационная сложность пропорциональна $1/c^2$, имеющихся данных попросту недостаточно. Однако использование множественных линейных аппроксимаций не решает проблему. Более правильное объяснение состоит в том, что корреляция c известна только с точностью до ошибки моделирования ϵ, если истинна гипотеза ①, а из уточненной гипотезы с неверным ключом следует, что ошибка моделирования ϵ должна быть меньше $2^{-n/2}$. Это будет играть важную роль в главе 8.

Для множественного линейного криптоанализа влияние гипотезы с неправильным ключом особенно важно. Если $|\Lambda|$ велико, то анализ часто можно упростить, потому что для неправильных ключей емкость близка к среднему значению $|\Lambda|/2^n$. Из обсуждения в разделе 6.1.2 следует, что если q мало по сравнению с $1/(C^{\mathsf{F}}_{v,u})^2$ и корреляции неизвестны, то

$$q = \sqrt{2|\Lambda|}\,\frac{\Phi^{-1}(P_{\mathsf{S}}) - \Phi^{-1}(P_{\mathsf{F}})}{\mathsf{Cap}(\Lambda) - |\Lambda|/2^n},$$

в предположении, что допущения из раздела 6.1.2 по-прежнему справедливы.

7.4. Оптимальное восстановление ключа

В главе 1 упоминался еще один подход к восстановлению ключа: алгоритм 1 Мацуи. Этот метод можно обобщить на задачу классификации: зная вектор эмпирических корреляций, найти наиболее вероятное значение битов ключа, которое определяет эти корреляции. Для данной задачи существует оптимальное решение, называемое байесовским классификатором.

Однако решением этой задачи классификации дело не ограничивается. Первая проблема заключается в том, что однозначно восстановить значение ключа может оказаться невозможно. Например, если корреляция равна $(-1)^{\kappa_1}/8\ (1 + (-1)^{\kappa_2}/2)(1 + (-1)^{\kappa_3}/2)$, то перемена мест κ_2 и κ_3 всегда приводит к равным правдоподобиям. Это наблюдение ведет ко второй проблеме: полезнее получить список возможных ключей, чем одного кандидата. Следовательно, существует компромисс между количеством классов ключей и вероятностью правильной классификации.

Мы здесь не обсуждаем оптимальный способ решения таких задач классификации, поскольку это завело бы нас в неисследованные дебри. Однако в завершение этого раздела стоит упомянуть, что подход к восстановлению ключа на основе «алгоритма 2» на самом деле больше похож на классификацию, чем на проверку гипотез. В разделах 7.2 и 7.3, как и в главе 4, предполагалось, что

процесс восстановления ключа можно рассматривать как форму множественной проверки гипотез. Однако при таком подходе множественные проверки гипотез в действительности не являются статистически независимыми. Зависимости между оценками корреляций для различных ключей проще учесть в схеме на основе классификации.

7.5. Историческая справка

Теория проверки простых гипотез Неймана и Пирсона впервые была применена к линейному криптоанализу в работе Бэне, Жюно и Воденэ. Их анализ применим к многомерному линейному криптоанализу и сравним с обсуждением в разделе 7.2.3.

Общая гипотеза рандомизации с неправильным ключом из раздела 7.3.3 была введена Богдановым и Тишхаузером. Термин «рандомизация с неправильным ключом» был предложен Харпесом, Крамером и Мэсси. Для одной линейной аппроксимации и выборки без возвращения, в частности, апостериорное распределение статистики критерия при условии истинности этой гипотезы обсуждалось в работе Ашура, Бейна и Рэймена (2020).

7.6. Литература

Ashur, Tomer, Tim Beyne, and Vincent Rijmen (Apr. 2020). «Revisiting the Wrong-Key-Randomization Hypothesis». In: *Journal of Cryptology* 33.2, pp. 567–594. doi: 10.1007/s00145-020-09343-2.

Baignères, Thomas, Pascal Junod, and Serge Vaudenay (Dec. 2004). «How Far Can We Go Beyond Linear Cryptanalysis?» In: *ASIACRYPT 2004*. Ed. by Pil Joong Lee. Vol. 3329. LNCS. Springer, Berlin, Heidelberg, pp. 432–450. doi: 10.1007/978-3-540-30539-2_31.

Bogdanov, Andrey and Elmar Tischhauser (Mar. 2014). «On the Wrong Key Randomisation and Key Equivalence Hypotheses in Matsui's Algorithm 2». In: *FSE 2013*. Ed. by Shiho Moriai. Vol. 8424. LNCS. Springer, Berlin, Heidelberg, pp. 19–38. doi: 10.1007/978-3-662-43933-3_2.

Harpes, Carlo, Gerhard G. Kramer, and James L. Massey (May 1995). «A Generalization of Linear Cryptanalysis and the Applicability of Matsui's Piling-Up Lemma». In: *EUROCRYPT'95*. Ed. by Louis C. Guillou and Jean-Jacques Quisquater. Vol. 921. LNCS. Springer, Berlin, Heidelberg, pp. 24–38. doi: 10.1007/3-540-49264-X_3.

Kullback, Solomon and Richard A. Leibler (1951). «On Information and Sufficiency». In: *The Annals of Mathematical Statistics* 22.1, pp. 79–86.

7.7. Упражнения

Упражнение 7.1

Докажите теорему 7.3 для случая равномерно распределенных случайных функций. Используйте многомерную центральную предельную теорему.

* Упражнение 7.2

Докажите теорему 7.3 для случая равномерно распределенных случайных перестановок.

Упражнение 7.3

Пусть Λ – множественная линейная аппроксимация, состоящая из линейных аппроксимаций с корреляциями $c_k \in \mathbb{R}^{|\Lambda|}$ для любого ключа k. Предположим, что априорное распределение $c_{\mathbf{k}}$ с равномерно распределенным случайным ключом $\mathbf{k}$ является многомерным нормальным распределением с нулевым средним и ковариационной матрицей Σ. Будем использовать предположения простой модели.

1. Покажите, что существует линейная замена переменных эмпирических корреляций, такая что статистика критерия логарифмического отношения правдоподобия с точностью до сдвига и масштабирования равна взвешенной сумме квадратов.
2. Покажите, что информационная сложность пропорциональна $1/\sqrt{\mathrm{Tr}\,\Sigma^{\intercal}\Sigma}$.

Глава 8

Аппроксимации с нулевой корреляцией

Традиционно в линейном криптоанализе используются линейные аппроксимации с атипично большой абсолютной корреляцией. Но в этой главе мы рассмотрим, как можно использовать аппроксимации с нулевой корреляцией. Этот вариант линейного криптоанализа называется *линейным криптоанализом с нулевой корреляцией*.

8.1. Идея

Грубо говоря, для атаки с восстановлением ключа в стиле алгоритма 2 Мацуи достаточно найти такое свойство внутренней части шифра, которое позволяет отличить правильные догадки от неправильных. В простой модели из главы 4 предполагается, что корреляции линейных аппроксимаций для неправильных ключей равны нулю. С точки зрения этой упрощенной модели, линейная аппроксимация с нулевой корреляцией бесполезна. Однако, как обсуждалось в главе 7, корреляции для неправильных ключей не в точности равны нулю при более точных гипотезах рандомизации с неправильным ключом, например в модели случайной перестановки.

Хотя из модели случайной перестановки следует, что линейные аппроксимации с нулевой корреляцией могли бы быть полезны *в принципе*, необходимо решить некоторые проблемы. Первая из них – нахождение линейных аппроксимаций с нулевой корреляцией. Трудность в том, что недостаточно, чтобы корреляция была малой, – она должна быть в точности равна нулю. Этой проблемой мы займемся в разделе 8.2. Еще один вопрос – являются ли аппроксимации с нулевой корреляцией достаточно «примечательными», чтобы принести пользу в качестве различающих свойств. Вероятность успеха всегда можно сделать близкой к единице, если использовать все возможные пары (открытый текст, шифртекст) для вычисления корреляции, но чтобы отфильтровать достаточное число неправильных ключей, должна быть также близка к нулю вероятность ложноположительного результата. В конце концов, согласно теореме 7.3, нуль по-прежнему является наиболее вероятным значением корреляции для случайной перестановки.

Причина, по которой низкие вероятности ложноположительного результата достижимы, связана с несколько противоречащим интуиции (по крайней мере для неспециалистов) свойством распределений вероятностей. По мере увели-

чения числа возможных исходов вероятность каждого отдельного исхода – даже самого вероятного – уменьшается. Но для вероятностей достаточно широких интервалов исходов это уже не так. Применяя данный факт к случаю линейного криптоанализа, мы видим, что для перестановки, выбранной равномерно случайным образом, большинство приближений будут иметь корреляцию, «близкую к нулю», но корреляция, в точности равная нулю, встречается редко.

Аппроксимацию с нулевой корреляцией можно использовать в качестве различающего свойства при условии, что наша оценка корреляции достаточно точна, чтобы различить нуль и «близко к нулю». Если доступна только одна аппроксимация с нулевой корреляцией, то это приводит к информационной сложности, близкой к 2^n для n-битовой функции. В разделах 8.3 и 8.4 обсуждаются методы уменьшения информационной сложности.

Наконец, заметим, что в принципе каждое значение корреляции (или даже диапазон значений) можно было бы использовать в качестве различителя, если корреляция известна достаточно точно. Нулевое значение является особым, поскольку часто проще показать, что некоторые линейные аппроксимации имеют нулевую корреляцию.

8.2. Нахождение аппроксимаций с нулевой корреляцией

Как показано в следствии 2.8, корреляцию линейной аппроксимации можно записать в виде суммы корреляций линейных следов. Хотя достаточно, чтобы была равна нулю сумма корреляций следов, почти все аппроксимации с нулевой корреляцией, описанные в литературе, обладают тем свойством, что все линейные следы имеют нулевую корреляцию.

Существование линейных аппроксимаций, таких что все следы имеют нулевую корреляцию, и тот факт, что некоторые из них легко найти, связаны с использованием функций с большим числом линейных разветвлений (см. определение 3.1) и, более общо, с использованием раундовых функций с простой структурой.

Если все линейные следы внутри линейной аппроксимации имеют нулевую корреляцию, то это можно проверить с помощью автоматизированных методов из главы 3. Достаточно искать линейные следы с ненулевой корреляцией, не пытаясь максимизировать их корреляцию: если ни одного решения не найдено, то линейная аппроксимация имеет нулевую корреляцию. К сожалению, этот подход не дает ответа на вопрос, какие линейные аппроксимации имеют нулевую корреляцию.

Ниже мы обсудим более информативный метод *потери посередине*, который часто можно применить вручную. Прежде чем обсуждать этот метод в общем виде, мы приведем пример для трех раундов демонстрационного шифра из раздела 1.1.

Пример 8.1 (потеря посередине). Рассмотрим три раунда демонстрационного шифра из раздела 1.1, представленных на рис. 8.1. Этот пример показывает, что линейная аппроксимация $(u, v) = (000000001, 000000001)$ имеет нулевую корреляцию. Чтобы установить это, мы рассмотрим все линейные аппроксимации с входной маской 000000001 в первой половине шифра и все линейные аппроксимации с выходной маской 000000001 во второй половине.

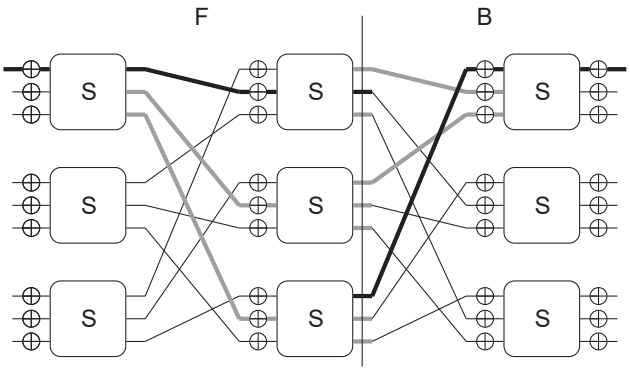


Рис. 8.1. Нахождение линейной аппроксимации с нулевой корреляцией методом потери посередине

Пусть $\mathsf{E}_k = \mathsf{B} \circ \mathsf{F}$, где F и B показаны на рис. 8.1. А именно F состоит из первого раунда и уровня S-блоков второго раунда, а B – из перестановки битов во втором раунде и уровне S-блоков третьего раунда. Корреляция (u, v) на E_k равна

$$C^{\mathsf{E}_k}_{v,u} = \sum_{w \in \mathbb{F}_2^9} C^{\mathsf{B}}_{v,w} C^{\mathsf{F}}_{w,u}.$$

Здесь, чтобы показать, что (u, v) – линейная аппроксимация с нулевой аппроксимацией, достаточно показать, что для всех следов (u, w, v) либо $C^{\mathsf{F}}_{w,u} = 0$, либо $C^{\mathsf{B}}_{w,u} = 0$.

Поскольку строка 001 матрицы C^{S} содержит ненулевые элементы только в столбцах $ab1$, где $a, b \in \mathbb{F}_2$, условие $C^{\mathsf{B}}_{w,u} \neq 0$ можно переписать в виде

$$w \in V = \{00100b00a\} \mid a,b \in \mathbb{F}_2\}.$$

Аналогично столбец 001 матрицы C^{S} содержит ненулевые элементы в строках $ab1$, где $a, b \in \mathbb{F}_2$. Следовательно, после первого раунда все маски, которые приводят к ненулевой корреляции, имеют вид $0b00a0010$. Из того, что столбец 010 матрицы C^{S} содержит ненулевые элементы только в строках $a1b$, где $a, b \in \mathbb{F}_2$, следует, что $C^{\mathsf{F}}_{w,u} \neq 0$ влечет за собой

$$w \in U = \{c \parallel a1b \mid a, b \in \mathbb{F}_2, c \in \mathbb{F}_2^6\}.$$

Если существует линейный след (u, w, v) для $\mathsf{B} \circ \mathsf{F}$ с ненулевой корреляцией, то $w \in U$ и $w \in V$, как было показано выше. Однако U и V не пересекаются, поэтому (u, v) – аппроксимация с нулевой корреляцией.

Проверьте, что то же рассуждение работает для любой маски v, в которой три средних бита равны нулю. Например, каждая из (000000001, 000000010), (000000001, 000000100) и (000000001, 000000110) – линейная аппроксимация с нулевой корреляцией.

В упражнении 8.1 вам будет предложено показать, что все линейные аппроксимации с входной маской 000001000 или 000001001 и выходной маской, в которой все биты, кроме первых трех, равны нулю, также имеют нулевую корреляцию. ▷

В общем случае метод потери посередине, проиллюстрированный в примере 8.1, работает следующим образом. Пусть $\mathsf{E}_k = \mathsf{B} \circ \mathsf{F}$ – шифр, где F и B соответствуют произвольному разложению на две части. Например, F могла бы состоять из первых r_F раундов шифра, а B – из остальных r_B раундов. Метод потери посередине приводит к аппроксимациям с нулевой корреляцией для $r_\mathsf{F} + r_\mathsf{B}$ раундов.

Начав с входной маски u, которая обычно имеет специальную структуру, например низкий вес Хэмминга, мы определяем множество выходных масок w такое, что (u, w) могла бы иметь ненулевую корреляцию на F. Как правило, это делается путем распространения масок вперед раунд за раундом, как в примере 8.1. Это дает множество масок U, такое что

$$U \supseteq \left\{ w \in \mathbb{F}_2^n \mid C_{w,u}^{\mathsf{F}} \neq 0 \right\}.$$

Аналогично, начав с выходной маски v, мы определяем множество масок V, такое что

$$V \supseteq \left\{ w \in \mathbb{F}_2^n \mid C_{v,w}^{\mathsf{B}} \neq 0 \right\}.$$

Для построения V маска v распространяется назад через B раунд за раундом. Множества U и V обычно описываются неявно, например с помощью паттернов или линейных уравнений, а не в терминах их элементов. Тогда корреляция линейной аппроксимации (u, v) имеет вид

$$C_{v,u}^{\mathsf{E}_k} = \sum_{w \in U \cap V} C_{v,w}^{\mathsf{B}} C_{w,u}^{\mathsf{F}}.$$

Следовательно, если $U \cap V$ пусто, то (u, v) – линейная аппроксимация с нулевой корреляцией.

Выбор аппроксимации (u, v) зависит от деталей функции. В типичном случае сама структура F и B предполагает один или несколько перспективных кандидатов.

8.3. Использование аппроксимаций с нулевой корреляцией

В этом разделе мы более подробно рассмотрим, как линейные аппроксимации с нулевой корреляцией можно использовать в качестве различителя. Анализ предполагает, что доступно достаточное количество пар (открытый текст, шифтекст) для вычисления точных корреляций аппроксимаций.

Если корреляции можно вычислить точно, то вероятность успеха равна единице. Первая цель этого раздела – вычислить соответствующую вероятность ложноположительного результата.

На первый взгляд, линейный криптоанализ с нулевой корреляцией может показаться непрактичным, потому что для вычисления точных корреляций, похоже, требуются все возможные пары (открытый текст, шифртекст). Вторая цель этого раздела – показать, что при наличии нескольких аппроксимаций часто бывает достаточно меньшего числа выбранных пар (открытый текст, шифртекст).

8.3.1. Одна аппроксимация

Аппроксимацию с нулевой корреляцией можно использовать для восстановления ключа, если воспользоваться проверкой статистических гипотез. Большинство методов из разделов 4.2 и 5.1 применимо без каких-либо изменений.

Как было объяснено в разделе 5.1, блочный шифр разбивается на внутреннюю и внешнюю части, состоящие из функций F_l и B_k. Аппроксимация с нулевой корреляцией применяется к внутренней части. Оценка корреляции для внутренней части вычисляется по формуле (5.1):

$$\widehat{c}_{k,l} = \frac{1}{q}\sum_{i=1}^{q}(-1)^{\mathsf{F}_l(x_i)+\mathsf{B}_k(y_i)}.$$

В отличие от обычного линейного криптоанализа, для вычисления $\hat{c}_{k,l}$ мы используем все возможные пары (открытый текст, шифртекст). Следовательно, «оценка» равна фактической корреляции. В частности, $\hat{c}_{k,l} = 0$, если k и l являются правильными ключами. Для неправильных ключей мы используем гипотезу рандомизации с неправильным ключом, основанную на модели случайной перестановки из раздела 7.3.3. В этой модели маловероятно, что $\hat{c}_{k,l} = 0$. Точная вероятность определяется следствием 8.2, которое вытекает из более общей теоремы 8.1. Эта вероятность является вероятностью ложноположительного результата проверки.

Теорема 8.1 (корреляция для случайной перестановки). *Пусть* F *– случайная равномерно распределенная перестановка на* $\mathbb{F}_2^n$. *Вероятность того, что линейная аппроксимация (u, v) перестановки* F *с* $u, v \neq 0$ *имеет корреляцию* $4w/2^n - 1$, *равна*

$$\Pr\left[C^{\mathsf{F}}_{v,u} = 4w/2^n - 1\right] = \frac{\binom{2^{n-1}}{w}^2}{\binom{2^n}{2^{n-1}}}$$

в предположении, что $w \leqslant 2^{n-1}$ *– неотрицательное число.*

Доказательство. Существует $2^n!$ перестановок на $\mathbb{F}_2^n$. Для доказательства результата достаточно подсчитать количество перестановок, таких что (u, v) имеет корреляцию $4w/2^n - 1$, где $0 \leqslant w \leqslant 2^{n-1}$. Для этого разобьем $\mathbb{F}_2^n$ на два подмножества: $\{x \in \mathbb{F}_2^n \mid u^{\mathsf{T}}x = 0\}$ и $\{x \in \mathbb{F}_2^n \mid u^{\mathsf{T}}x = 1\}$. Это приводит к следующему распределению входных значений (как в доказательстве теоремы 1.1):

	$u^{\mathsf{T}}x = 0$	$u^{\mathsf{T}}x = 1$
$v^{\mathsf{T}}\mathsf{F}(x) = 0$	w	$2^{n-1} - w$
$v^{\mathsf{T}}\mathsf{F}(x) = 1$	$2^{n-1} - w$	w

Существует 2^{n-1} значений x, таких что $u^{\mathsf{T}}x = 0$. Что касается образов $\mathsf{F}(x)$, значения w должны принадлежать множеству $\{y \in \mathbb{F}_2^n \mid v^{\mathsf{T}}y = 0\}$, а значения $2^{n-1} - w$ – его дополнению. Следовательно, число способов выбрать образы равно

$$\binom{2^{n-1}}{w}\binom{2^{n-1}}{2^{n-1}-w} = \binom{2^{n-1}}{w}^2.$$

Существует $2^{n-1}!$ способов сопоставить эти образы входам x, поэтому число сопоставлений образов $\mathsf{F}(x)$ входам x, таким что $u^{\mathsf{T}}x = 0$, равно

$$2^{n-1}!\binom{2^{n-1}}{w}^2.$$

Что касается сопоставлений образов $\mathsf{F}(x)$ входам x, таким что $u^{\mathsf{T}}x = 1$, то множество $2^{n-1} - w$ образов, принадлежащих $\{y \in \mathbb{F}_2^n \mid v^{\mathsf{T}}y = 0\}$, должно быть дополнением множества значений w, которые уже были выбраны для случая $u^{\mathsf{T}}x = 0$. Аналогично множество значений w, принадлежащих $\{y \in \mathbb{F}_2^n \mid v^{\mathsf{T}}y = 0\}$, уже было определено. Следовательно, поскольку существует $2^{n-1}!$ способов сопоставить образы входам, количество перестановок с корреляцией $4w/2^n - 1$ в точности равно

$$\left(2^{n-1}!\right)^2\binom{2^{n-1}}{w}^2.$$

Деление на общее количество перестановок дает следующую вероятность:

$$\frac{\left(2^{n-1}!\right)^2\binom{2^{n-1}}{w}^2}{2^n!} = \frac{\binom{2^{n-1}}{w}^2}{\binom{2^n}{2^{n-1}}}.$$

На этом доказательство завершается. □

Следствие 8.2 (нулевая корреляция для случайной перестановки). *Пусть* F – *случайная равномерно распределенная перестановка на* $\mathbb{F}_2^n$. *Вероятность того, что линейная аппроксимация (u, v) перестановки* F *с* $u, v \neq 0$ *имеет нулевую корреляцию, равна*

$$\Pr\left[C_{v,u}^{\mathsf{F}} = 0\right] = \frac{\binom{2^{n-1}}{2^{n-2}}^2}{\binom{2^n}{2^{n-1}}} = 2\sqrt{\frac{2}{\pi}}\,2^{-n/2} + \mathcal{O}\left(2^{-3n/2}\right),$$

когда $n \to \infty$.

Доказательство. Результат следует из теоремы 8.1, если положить $w = 2^{n-2}$. Асимптотическое равенство вытекает из следующей оценки:

$$\binom{2N}{N} = \frac{2^{2N}}{\sqrt{\pi}}\left(N^{-1/2} + \mathcal{O}(N^{-3/2})\right),$$

которая является следствием аппроксимации факториала по формуле Стирлинга. □

Из следствия 8.2 вытекает, что если существует K возможных ключей, то приблизительно $P_{\mathsf{F}}K \approx 0.8 \times K/2^{n/2}$ из них остаются после фильтрации. Хотя

встречаются случаи, когда ключей больше, чем $2^{n/2}$, обычно это не составляет проблемы, потому что часто доступно более одной линейной аппроксимации с нулевой корреляцией.

Недостаток линейного криптоанализа с нулевой корреляцией заключается в том, что поскольку для вычисления корреляции используются все возможные пары (открытый текст, шифртекст), информационная сложность равна 2^n. Если внешняя часть шифра состоит только из одного или более конечных раундов, т. е. $\mathsf{F}_i(x) = u^\mathsf{T}x$, где u – входная маска, то, в силу упражнения 1.7, достаточно зашифровать все входы, принадлежащие множеству $\{x \in \mathbb{F}_2^n \mid u^\mathsf{T}x = 0\}$. Однако информационная сложность 2^{n-1} в большинстве случаев все еще непрактична. В разделе 8.3.2 показано, что если доступно несколько аппроксимаций с нулевой корреляцией, то информационную сложность можно уменьшить.

8.3.2. Несколько аппроксимаций

Пусть Λ – многомерная линейная аппроксимация. Если емкость Λ равна нулю, то, в силу следствия 6.6,

$$\Pr\left[(\mathbf{x}, \mathsf{F}(\mathbf{x})) \equiv (s,t) \bmod \Lambda^\perp\right] = \frac{1}{|\Lambda|} \sum_{(u,v)\in\Lambda} (-1)^{u^\mathsf{T}s+v^\mathsf{T}t} C_{v,u}^{\mathsf{F}} = \frac{1}{|\Lambda|}.$$

Иными словами, если $\mathbf{x}$ – случайная равномерно распределенная величина, то таковой является и $(\mathbf{x}, \mathsf{F}(\mathbf{x})) \bmod \Lambda^\perp$. Обратное тоже верно: если $(\mathbf{x}, \mathsf{F}(\mathbf{x})) \bmod \Lambda^\perp$ распределена равномерно, то ненулевые пары в Λ являются линейными аппроксимациями с нулевой корреляцией.

Это дает альтернативное описание многомерного линейного криптоанализа с нулевой корреляцией, но не уменьшает информационную сложность. Однако если $\Lambda = \Lambda_{\text{in}} \oplus \Lambda_{\text{out}}$, где $\Lambda_{\text{in}} \subseteq \mathbb{F}_2^n$ и $\Lambda_{\text{out}} \subseteq \mathbb{F}_2^n$, то, в силу следствия 6.6,

$$\Pr\left[\mathsf{F}(\mathbf{x}) \equiv t \bmod \Lambda_{\text{out}}^\perp\right] = \frac{1}{|\Lambda_{\text{out}}|} \sum_{(u,v)\in\Lambda} (-1)^{u^\mathsf{T}s+v^\mathsf{T}t} C_{v,u}^{\mathsf{F}} = \frac{1}{|\Lambda_{\text{out}}|}$$

для случайной величины $\mathbf{x}$, равномерно распределенной на $s + \Lambda_{\text{in}}^\perp$. Этот результат можно использовать для уменьшения информационной сложности. А именно достаточно зашифровать множество вида $s + \Lambda_{\text{in}}^\perp$. Тогда для различения нужно проверить, что множество входов x, таких что $\mathsf{F}(x) \equiv t \bmod \Lambda_{\text{out}}^\perp$, одинаково для всех значений $t \in \mathbb{F}_2^n/\Lambda_{\text{out}}^\perp$. Следовательно, информационная сложность составляет всего $2^n/|\Lambda_{\text{in}}|$.

Пример 8.2. Пусть $\Lambda = \Lambda_{\text{in}} \oplus \Lambda_{\text{out}}$, где $\Lambda_{\text{in}} = \text{Span}\{000000001, 000001000\}$ и $\Lambda_{\text{out}} = \text{Span}\{000000001, 000000010, 000000100\}$. В примере 8.1 и в упражнении 8.1 показано, что Λ является многомерной линейной аппроксимацией с нулевой корреляцией для трех раундов демонстрационного шифра. Следовательно, если $\mathbf{x}$ – случайная величина, равномерно распределенная на $s + \Lambda_{\text{in}}^\perp$, то

$$\Pr\left[\mathsf{E}_k(\mathbf{x}) \equiv t \bmod \Lambda_{\text{out}}^\perp\right] = \frac{1}{8},$$

где E_k – три раунда демонстрационного шифра.

Векторное пространство $\Lambda_{\text{in}}^{\perp}$ состоит из всех значений $x = (x_8, \ldots, x_0)$, таких что x_0 и x_3 равны нулю. Следовательно, $\mathbf{x}$ – случайная величина, равномерно распределенная на множестве открытых текстов, в которых эти два бита постоянны.

Аналогично векторное пространство $\Lambda_{\text{out}}^{\perp}$ состоит из всех значений y, таких что первые три бита y равны нулю. Следовательно, первые три бита $\mathsf{E}_k(x)$ однозначно представляют $\mathsf{E}_k(x) \bmod \Lambda_{\text{out}}^{\perp}$. ▷

Тот факт, что $\mathsf{F}(\mathbf{x}) \bmod \Lambda^{\perp}$ – случайная равномерно распределенная величина, называют также свойством *насыщения*. В главе 9 мы вернемся к этим свойствам в контексте атак с насыщением.

Для вычисления вероятности ложноположительного результата многомерного различителя с нулевой корреляцией модель случайной перестановки используется в качестве гипотезы рандомизации с неправильным ключом. Результатом является показанный ниже вариант следствия 8.2. Доказательство основано на функции массы вероятности многомерного гипергеометрического распределения, которое можно вывести с помощью рассуждения, аналогичного доказательству теоремы 8.1. В упражнении 8.5 вам будет предложено дать полное доказательство.

Теорема 8.3. *Пусть* F *– случайная равномерно распределенная перестановка на* $\mathbb{F}_2^n$*, и пусть* $\Lambda = \Lambda_{\text{in}} \oplus \Lambda_{\text{out}}$ *– многомерная линейная аппроксимация* F *с* $|\Lambda| \leqslant 2^n$*. Вероятность того, что* $\Pr_{\mathbf{x}}[\mathsf{F}(\mathbf{x}) \equiv t \bmod \Lambda_{\text{out}}^{\perp}] = 1/|\Lambda_{\text{out}}|$ *для всех* $t \in \mathbb{F}_2^n$ *и случайной величины* $\mathbf{x}$*, равномерно распределенной на* $s + \Lambda_{\text{in}}^{\perp}$ *для некоторого* $s \in \mathbb{F}_2^n$*, равна*

$$\binom{|\Lambda_{\text{out}}^{\perp}|}{2^n/|\Lambda|}^{|\Lambda_{\text{out}}|} \Big/ \binom{2^n}{|\Lambda_{\text{in}}^{\perp}|}.$$

8.4. Статистический подход

В разделе 8.3 предполагалось, что корреляции или распределения вероятностей (для многомерных аппроксимаций) должны быть вычислены точно, чтобы можно было использовать аппроксимации с нулевой корреляцией в качестве различителей. Это требование можно ослабить, но ценой уменьшения вероятности успеха.

Как и в обыкновенном линейном криптоанализе, корреляции или – в многомерном случае – распределения вероятностей можно оценить с помощью случайной выборки пар (открытый текст, шифртекст). Для правильного ключа эмпирические корреляции будут равны нулю в среднем с дисперсией $1/q$. Для неправильных ключей среднее эмпирических корреляций близко к нулю, но не равно нулю в точности, а их дисперсия приближенно равна $1/q$. Эти два распределения одинаковы, как при анализе (множественного) линейного криптоанализа в простой модели с неизвестными корреляциями, с тем отличием, что роли правильного и неправильного ключа поменялись местами.

Средняя емкость многомерной линейной аппроксимации Λ случайной равномерно распределенной перестановки F на $\mathbb{F}_2^n$ равна

$$\mathbb{E}_{\mathsf{F}}\,\mathsf{Cap}(\Lambda) = \sum_{\substack{(u,v)\in\Lambda \\ (u,v)\neq(0,0)}} \mathbb{E}_{\mathsf{F}}\left(C_{v,u}^{\mathsf{F}}\right)^2 = |\Lambda|/2^n.$$

Не приводя строгого доказательства, заметим, что $\mathsf{Cap}(\Lambda)$ с высокой вероятностью близка к своему среднему значению. Следовательно, в модели случайной перестановки емкость линейной аппроксимации с нулевой корреляцией приближенно одинакова для всех неправильных ключей. В силу анализа из разделов 6.1.2 и 7.3.3, информационная сложность различителя, основанного на многомерной линейной аппроксимации с нулевой корреляцией, равна

$$q = \sqrt{2|\Lambda|}\,\frac{\Phi^{-1}(P_{\mathsf{S}}) - \Phi^{-1}(P_{\mathsf{F}})}{\mathsf{Cap}(\Lambda)} \approx \left(\Phi^{-1}(P_{\mathsf{S}}) - \Phi^{-1}(P_{\mathsf{F}})\right)\frac{2^{n+\frac{1}{2}}}{\sqrt{|\Lambda|}}$$

для $P_{\mathsf{S}} \geqslant P_{\mathsf{F}}$. Второе равенство опирается на аппроксимацию $\mathsf{Cap}(\Lambda) \approx |\Lambda|/2^n$. Можно показать, что это оптимально в смысле среднего случая из раздела 7.3.1, при условии что в качестве гипотезы рандомизации с неправильным ключом используется модель случайной перестановки.

Как было объяснено в разделе 6.2.3, если $\Lambda = \Lambda_{\mathrm{in}} \oplus \Lambda_{\mathrm{out}}$, то открытые тексты можно выбирать из смежного класса $\Lambda_{\mathrm{in}}^{\perp}$, чтобы уменьшить информационную сложность. Точнее, при $P_{\mathsf{S}} \geqslant P_{\mathsf{F}}$ информационная сложность становится равной

$$q = \sqrt{2|\Lambda_{\mathrm{out}}|}\,\frac{\Phi^{-1}(P_{\mathsf{S}}) - \Phi^{-1}(P_{\mathsf{F}})}{\mathsf{Cap}(\Lambda)} \approx \left(\Phi^{-1}(P_{\mathsf{S}}) - \Phi^{-1}(P_{\mathsf{F}})\right)\frac{2^{n+\frac{1}{2}}}{|\Lambda_{\mathrm{in}}|\sqrt{|\Lambda_{\mathrm{out}}|}}.$$

В самом деле, средняя емкость остается без изменения, но число аппроксимаций снижается до $|\Lambda_{\mathrm{out}}|$. Интуитивно понятно, что информационная сложность дополнительно уменьшается в $\sqrt{|\Lambda_{\mathrm{in}}|}$ раз сверх улучшения, описанного в разделе 8.3.2.

8.5. Историческая справка

Линейный криптоанализ с нулевой корреляцией был введен в рассмотрение Богдановым и Рэйменом. Они использовали метод потери посередине, который мы обсуждали в разделе 8.2, чтобы найти линейные аппроксимации с нулевой корреляцией.

Статистический подход, описанный в разделе 8.4, который основан на нескольких линейных аппроксимациях с нулевой корреляцией, был предложен Богдановым и Ванем. Улучшение за счет выбранного открытого текста впервые использовали Богданов, Леандр, Нюберг и Вань.

8.6. Литература

Bogdanov, Andrey et al. (Dec. 2012). «Integral and Multidimensional Linear Distinguishers with Correlation Zero». In: *ASIACRYPT 2012*. Ed. by Xiaoyun Wang and Kazue Sako. Vol. 7658. LNCS. Springer, Berlin, Heidelberg, pp. 244–261. doi: 10.1007/978-3-642-34961-4_16.

Bogdanov, Andrey and Vincent Rijmen (2014). «Linear Hulls with Correlation Zero and Linear Cryptanalysis of Block Ciphers». In: *DCC* 70.3, pp. 369–383. doi: 10.1007/s10623-012-9697-z.

Bogdanov, Andrey and Meiqin Wang (Mar. 2012). «Zero Correlation Linear Cryptanalysis with Reduced Data Complexity». In: *FSE 2012*. Ed. by Anne Canteaut. Vol. 7549. LNCS. Springer, Berlin, Heidelberg, pp. 29–48. doi: 10.1007/978-3-642-34047-5_3.

8.7. Упражнения

Упражнение 8.1

Покажите, что для всех $u \in \mathbb{F}_2^3$ (000001000, 000000||u) и (000001001, 000000||u) являются линейными аппроксимациями с нулевой корреляцией для трех раундов демонстрационного шифра.

Упражнение 8.2

Пусть $\mathsf{E}_k : \mathbb{F}^6_2 \to \mathbb{F}^6_2$ – конструкция на рис. 8.2 (см. также рис. 2.4). Найдите нетривиальную линейную аппроксимацию с нулевой корреляцией для всех значений k.

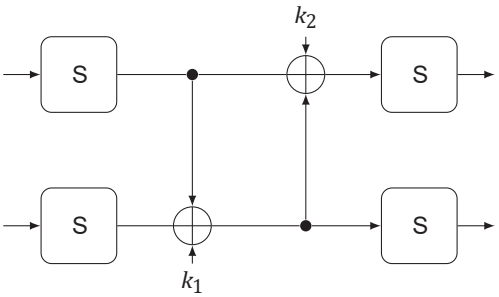


Рис. 8.2. Конструкция с четырьмя S-блоками

Упражнение 8.3

Рассмотрим пятираундовый шифр Фейстеля; первые два раунда показаны на рис. 8.3. Покажите, что если $\mathsf{F}_1, \ldots, \mathsf{F}_5$ – перестановки, то (0||u, u||0) – аппроксимация с нулевой корреляцией для всех ненулевых значений u.

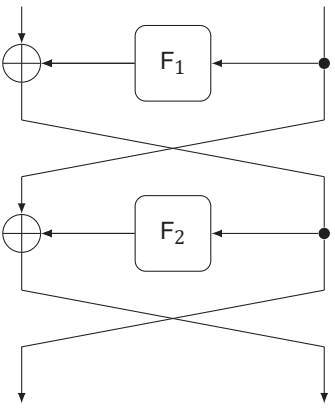


Рис. 8.3. Два раунда сети Фейстеля

Упражнение 8.4

Конструкцию «шифрование–перемешивание–шифрование» (рис. 8.4) можно использовать для построения блочного шифра, основанного на пяти функциях с половинным размером блока. Ваша задача – отличить выход конструкции «шифрование–перемешивание–шифрование» от выхода случайной равномерно распределенной перестановки $2n$ бит. На рис. 8.4 E_1, E_2, E_3 E_4 – блочные шифры с размером блока n бит и секретным ключом.

1. Предположим, что $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^n$– перестановка. Найдите многомерную аппроксимацию с нулевой корреляцией конструкции «шифрование–перемешивание–шифрование», содержащую 2^{2n} линейных аппроксимаций.
2. Основываясь на ответе на предыдущий вопрос, найдите время и число выбранных открытых текстов, необходимые, чтобы отличить «шифрование–перемешивание–шифрование» от случайной равномерно распределенной перестановки $2n$ бит?

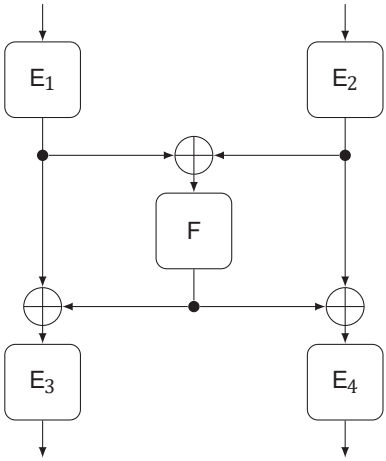


Рис. 8.4. Конструкция «шифрование–перемешивание–шифрование»

Упражнение 8.5

Пусть F – случайная равномерно распределенная перестановка на $\mathbb{F}_2^n$, и пусть $\Lambda = \Lambda_{\text{in}} \oplus \Lambda_{\text{out}}$ – многомерная линейная аппроксимация F.

1. Докажите теорему 8.3.
2. Покажите, что если Λ – многомерная линейная аппроксимация с нулевой корреляцией (для произвольной функции), то $|\Lambda| < 2^n$. Следовательно, это условие всегда выполняется при применении теоремы 8.3 на практике.

* Упражнение 8.6

Следствие 8.2 и теорема 8.3 сформулированы для случайных перестановок, но линейный криптоанализ с нулевой корреляцией применим также к функциям, не являющимся обратимыми.

1. Докажите аналоги следствия 8.2 и теоремы 8.3, когда F – случайная равномерно распределенная функция $\mathbb{F}_2^n \to \mathbb{F}_2^n$.
2. Объясните, почему и как условие $u, v \neq 0$ в следствии 8.2 можно ослабить, когда F – случайная равномерно распределенная функция.

Упражнение 8.7

Цель этого упражнения – найти линейные атаки с нулевой корреляцией на демонстрационный шифр типа Rijndael из раздела 3.2.1 при небольшом числе раундов.

1. Найдите линейную аппроксимацию с нулевой корреляцией для четырех раундов.
2. Обобщите свою линейную аппроксимацию с нулевой корреляцией на многомерную аппроксимацию с нулевой корреляцией, которая требует как можно меньше данных. Какова получившаяся в итоге информационная сложность?
3. Обобщите свой различитель на атаку с восстановлением ключа на пять раундов и оцените временную и информационную сложность.

* Упражнение 8.8

Найдите линейную аппроксимацию с нулевой корреляцией для десяти раундов демонстрационного шифра типа Rijndael из раздела 3.2.1.

Упражнение 8.9

В этом упражнении исследуется *корреляционная атака на инвариант различия ключей*. Пусть $\mathsf{E}_k : \mathbb{F}_2^n \to \mathbb{F}_2^n$ – блочный шифр с чередованием ключа с раундовыми ключами $k = (k_1, \ldots, k_r)$. То есть $\mathsf{E}_k = \mathsf{R}_{k_r} \circ \cdots \circ \mathsf{R}_{k_2} \circ \mathsf{R}_{k_1}$, где $\mathsf{R}_{k_i}(x) = \mathsf{R}(x) + k_i$.

1. Обозначим $d = d_1 \| \ldots \| d_r \in \mathbb{F}_2^{nr}$ – различие между двумя раундовыми ключами, и пусть

$$\Lambda = \left\{ (u_2, u_3, \ldots, u_{r+1}) \mid u_2, \ldots, u_r \in \mathbb{F}_2^n \text{ и } \prod_{i=1}^{r} C^{\mathsf{R}}_{u_{i+1}, u_i} \neq 0 \right\}$$

– множество линейных следов с ненулевой корреляцией для аппроксимации (u_1, u_{r+1}). Приведите достаточные условия, которым должны удовлетворять d и Λ, чтобы гарантированно выполнялось равенство

$$C^{\mathsf{E}_k}_{u_{r+1}, u_1} = C^{\mathsf{E}_{k+d}}_{u_{r+1}, u_1},$$

где $k + d = (k_1 + d_1, k_2 + d_2, \ldots, k_r + d_r)$.

2. Найдите линейную аппроксимацию демонстрационного шифра из раздела 1, имеющую корреляцию инварианта различия ключей, но ненулевую корреляцию. Попытайтесь максимизировать число раундов.

Глава 9

Различные обобщения

Основные обобщения линейного криптоанализа были представлены в предыдущих главах; это множественный, многомерный и линейный криптоанализ с нулевой корреляцией. Однако это далеко не все расширения, предложенные в литературе. В данной главе представлен обзор некоторых наиболее важных предложений.

Большинство обсуждаемых ниже обобщений линейного криптоанализа являются отчасти гипотетическими: они показывают, как определенные комбинаторные свойства могут быть использованы для атаки на криптографические примитивы, но не дают точного рецепта анализа или нахождения этих свойств. В главе 11 мы вернемся к этому вопросу.

9.1. Точные свойства

Корреляции линейных аппроксимаций обычно известны только с некоторой погрешностью аппроксимации, поскольку невозможно учесть все линейные следы. Как обсуждалось в главе 8, линейный криптоанализ с нулевой корреляцией отличается тем, что в нем используется тот факт, что корреляция линейной аппроксимации в точности равна нулю. Оказывается, что большинство широко применимых обобщений линейного криптоанализа (кроме множественного и многомерного) основаны на подобных «точных» свойствах.

9.1.1. Атаки с насыщением

Как было объяснено в главе 8, если $\Lambda = \Lambda_{\text{in}} \oplus \Lambda_{\text{out}}$, где $\Lambda_{\text{in}} \subseteq \mathbb{F}_2^n$ и $\Lambda_{\text{out}} \subseteq \mathbb{F}_2^m$ – многомерная линейная аппроксимация с нулевой корреляцией функции $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$, то

$$\Pr\left[\mathsf{F}(\mathbf{x}) \equiv t \bmod \Lambda_{\text{out}}^{\perp}\right] = \frac{1}{|\Lambda_{\text{out}}|},$$

где $\mathbf{x}$ – случайная равномерно распределенная величина на смежном классе $\Lambda_{\text{in}}^{\perp}$ и для любого $t \in \mathbb{F}_2^m$. Иными словами, если все элементы смежного класса $\Lambda_{\text{in}}^{\perp}$ зашифрованы, то любое приведение шифртекста по модулю $\Lambda_{\text{out}}^{\perp}$ имеет место равное число раз. В главе 8 уже упоминалось, что это называется свойством насыщения.

Свойства насыщения иногда можно найти, анализируя распространение значений, а не линейные следы. Фактически именно так были найдены первые свойства насыщения еще до открытия линейного криптоанализа с нуле-

вой корреляцией. Однако такой анализ иногда выявляет также другие свойства шифртекста. Например, некоторые биты шифртекста могут оставаться неизменными, когда часть открытого текста насыщена. Это иллюстрируется в следующем примере.

Пример 9.1. На рис. 9.1 множество открытых текстов с одной насыщенной ячейкой распространяется через пять раундов демонстрационного шифра типа Rijndael из раздела 3.2.1. Точнее, входное множество состоит из восьми открытых текстов, так что все ячейки, кроме первой, постоянны, а первая ячейка принимает все возможные значения из $\mathbb{F}_2^3$ по одному разу. Нетрудно распространить это множество через первые несколько раундов шифра. Для этого пометим ячейку буквой A, если она принимает каждое 3-битовое значение одинаковое число раз (насыщена), буквой C, если она постоянна, и знаком «?» в противном случае. После пяти раундов гарантируется, что одна из ячеек состояния будет постоянной.

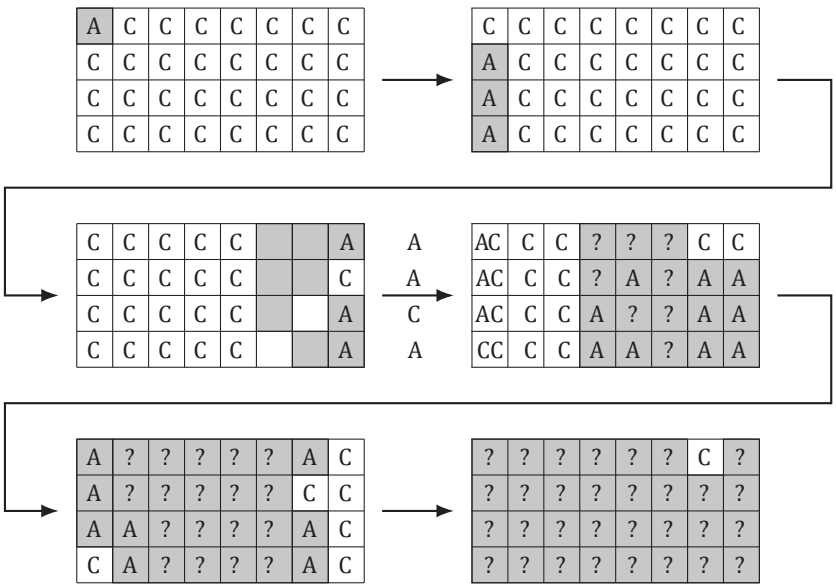


Рис. 9.1. Свойство для пяти раундов демонстрационного шифра типа Rijndael

В силу следствия 6.6, описанное выше свойство эквивалентно многомерной линейной аппроксимации $\Lambda_{\text{in}} \oplus \Lambda_{\text{out}}$. В частности,

$$\Lambda_{\text{in}} = \{x \| 00\cdots0 \mid x \in \mathbb{F}_2^3\}^{\perp} = \{000 \| x \mid x \in \mathbb{F}_2^{93}\}.$$

Кроме того, множество выходных масок Λ_{out} состоит из всех масок, равных нулю всюду, кроме постоянной ячейки. Многомерная аппроксимация $\Lambda = \Lambda_{\text{in}} \oplus \Lambda_{\text{out}}$ не является аппроксимацией с нулевой корреляцией. Вместо этого следствие 6.6 дает:

$$\sum_{(u,v)\in\Lambda} (-1)^{u^{\mathsf{T}}s+v^{\mathsf{T}}t} C^{\mathsf{F}}_{v,u} = |\Lambda_{\mathsf{out}}|$$

для любого s и конкретного t, зависящего от s. Таким образом, свойство на рис. 9.1 соответствует большому множеству линейных аппроксимаций, которые в сумме дают очень большое значение. На первый взгляд, это может показаться удивительным, но к тому же выводу можно прийти, рассуждая о линейных следах. ▷

Атаки с насыщением получили обобщение в двух направлениях. Первое, называемое *статистическими атаками с насыщением*, будет рассмотрено в разделе 9.2.1. Второе направление – *интегральный криптоанализ* – является самостоятельной областью исследований наряду с линейным криптоанализом. В этой книге он не обсуждается.

9.1.2. Инвариантные подпространства

Инвариантным подпространством функции $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^n$ называется аффинное подпространство $a + V$ пространства $\mathbb{F}_2^n$ такое, что $\mathsf{F}(a + V) \subseteq a + V$. Если F – перестановка, то $\mathsf{F}(a + V) = a + V$. Если инвариантное подпространство существует, то оно сразу же приводит к распознавателю с выбранным открытым текстом, для которого вероятность успеха близка к 1, а вероятность ложноположительного результата – к нулю. Инвариантные подпространства можно использовать для атак с восстановлением ключа, выразив условие, что частично дешифрированный шифртекст является элементом $a + V$, в виде системы уравнений. Однако это работает, только если инвариант не имеет места для неправильно угаданных ключей.

Исторически считалось, что инвариантные подпространства – следствие «очевидных» симметрий в шифре. Это иллюстрируется следующим примером.

Пример 9.2. Если игнорировать раундовые константы и ключи, то инвариантное подпространство существует для любого числа раундов демонстрационного шифра типа Rijndael из раздела 3.2.1:

x_1	y_1	x_1	y_1	x_1	y_1	x_1	y_1
y_1	x_1	y_1	x_1	y_1	x_1	y_1	x_1
x_2	y_2	x_2	y_2	x_2	y_2	x_2	y_2
y_2	x_2	y_2	x_2	y_2	x_2	y_2	x_2

Здесь x_1, x_2, y_1 и y_2 – произвольные значения, взятые из $\mathbb{F}_2^3$. Для большинства ключей это подпространство не сохраняется после сложения с раундовым ключом. Тем не менее существует множество, содержащее 2^{12} из 2^{96} ключей, для которого это подпространство инвариантно. Однако прибавление раундовых констант, не обладающих такой же симметрией, гарантированно препятствует продолжению этого свойства на несколько раундов. ▷

Инвариантные подпространства, которые действительно зависят от деталей S-блока и линейного уровня, были найдены для нескольких шифров, начиная с 2011 года. Эти подпространства обычно инвариантны только для подмножества ключей, которые поэтому называются *слабыми ключами*. Чем больше слабых ключей, тем выше вероятность успеха атаки. Эти «более тонкие» инвариантные подпространства обычно отыскиваются путем пораундового анализа, хотя и не все они могут быть найдены таким способом.

Пример 9.3. У демонстрационного шифра типа Rijndael из раздела 3.2.1 имеется инвариантное подпространство для 2^{32} из 2^{96} ключей. А именно пусть $U = \{000,111\}$. Так как $S(000) = 111$ и $S(111) = 000$, имеет место равенство $S(V = V$. Далее

$$\begin{bmatrix} 0 & I & I & I \\ I & 0 & I & I \\ I & I & 0 & I \\ I & I & I & 0 \end{bmatrix} \begin{bmatrix} 111 \\ 000 \\ 000 \\ 000 \end{bmatrix} = \begin{bmatrix} 000 \\ 111 \\ 111 \\ 111 \end{bmatrix}, \quad \begin{bmatrix} 0 & I & I & I \\ I & 0 & I & I \\ I & I & 0 & I \\ I & I & I & 0 \end{bmatrix} \begin{bmatrix} 000 \\ 000 \\ 111 \\ 111 \end{bmatrix} = \begin{bmatrix} 000 \\ 000 \\ 111 \\ 111 \end{bmatrix},$$

$$\begin{bmatrix} 0 & I & I & I \\ I & 0 & I & I \\ I & I & 0 & I \\ I & I & I & 0 \end{bmatrix} \begin{bmatrix} 000 \\ 111 \\ 111 \\ 111 \end{bmatrix} = \begin{bmatrix} 111 \\ 000 \\ 000 \\ 000 \end{bmatrix}, \quad \begin{bmatrix} 0 & I & I & I \\ I & 0 & I & I \\ I & I & 0 & I \\ I & I & I & 0 \end{bmatrix} \begin{bmatrix} 111 \\ 111 \\ 111 \\ 111 \end{bmatrix} = \begin{bmatrix} 111 \\ 111 \\ 111 \\ 111 \end{bmatrix}.$$

В силу симметрии матрицы M, отсюда следует, что $M \bigoplus_{i=1}^{4} U = \bigoplus_{i=1}^{4} U$. Следовательно, $V = \bigoplus_{i=1}^{32} U$ – инвариантное подпространство для `MixColumns` ∘ `ShiftRows` ∘ `SubCells`. Так как каждая ячейка раундовых констант равна 000 или 111, векторное пространство V также является инвариантным подпространством для сложения с раундовой константой. Однако V является инвариантным для шага сложения с ключом, только если все ячейки раундовых ключей равны 000 или 111. Таким образом, всего существует 2^{32} слабых ключей. ▷

Инвариантное подпространство в примере 9.3 не было найдено систематическим способом. Более того, даже если было возможно перечислить все инвариантные подпространства `SubCells`, `ShiftRows` и `MixColumns`, могли бы остаться другие инвариантные подпространства, кроме являющихся общими для `SubCells`, `ShiftRows` и `MixColumns`.

Если инвариантное подпространство достаточно велико и если количество слабых ключей не слишком мало, то практически возможно найти его, применяя подход на основе черного ящика. Пусть $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^n$ – перестановка. Чтобы найти наименьшее инвариантное подпространство F, которое содержит значение x, можно воспользоваться алгоритмом 9.1. Если F имеет нетривиальное инвариантное подпространство, то повторение алгоритма 9.1 со случайно выбранными x рано или поздно вернет это подпространство.

Алгоритм 9.1. Нахождение наименьшего инвариантного подпространства перестановки F

Вход:
Перестановка $\mathsf{F}\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$
Вектор $x \in \mathbb{F}_2^n$
Выход: наименьшее инвариантное подпространство F, содержащее x.
1: ▷ В алгоритме U может быть компактно представлено своим базисом
2: $U \leftarrow \{0\}$
3: **repeat**
4: $V \leftarrow U$
5: $U \leftarrow \mathrm{Span}\{x + \mathsf{F}(x + z) \mid z \in V\}$
6: **until** $U = V$
7: **return** $x + U$

Если F имеет инвариантное подпространство $a + V$ размерности d, то вероятность, что случайно выбранное x окажется в $a + V$, равна $2^d/2^n$. Следовательно, после повторения 2^{n-d} раз алгоритм 9.1 найдет (в среднем) подпространство $a + V$. В алгоритме 9.1 подпространства U и V могут быть компактно представлены своими базисами. Следовательно, общая временная сложность равна $O(n^3 2^{n-d})$.

9.1.3. Нелинейные инварианты

Нелинейным инвариантом функции $\mathsf{F}\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$ называется функция $f : \mathbb{F}_2^n \to \mathbb{F}_2$, такая что существует константа $b \in \mathbb{F}_2$, такая что для всех $x \in \mathbb{F}_2^n$

$$f(\mathsf{F}(x)) = f(x) + b.$$

Иными словами, $\mathsf{C}(f \circ \mathsf{F}, f) = (-1)^b$. Если F – блочный шифр, то b может зависеть от ключа.

Каждое инвариантное подпространство порождает нелинейный инвариант с $b = 0$. Действительно, если $a + V$ – инвариантное подпространство, то положим $f(x) = 1$, если $x \in a + V$, и $f(x) = 0$ в противном случае. В общем случае представлять себе нелинейный инвариант f можно также как множество $S = \{x \in \mathbb{F}_2^n \mid f(x) = 1\}$, такое что либо $\mathsf{F}(S) \subseteq S$ и $\mathsf{F}(\mathbb{F}_2^n \setminus S) \subseteq \mathbb{F}_2^n \setminus S$, либо $\mathsf{F}(S) \subseteq \mathbb{F}_2^n \setminus S$ и $\mathsf{F}(\mathbb{F}_2^n \setminus S) \subseteq S$. Первый случай соответствует $b = 0$, второй – $b = 1$.

Следующая теорема показывает, что для некоторых линейных функций, представленных блочной матрицей с единичными блоками, найти нелинейные инварианты легко. *Степенью* булевой функции называется степень ее полиномиального представления, а квадратичной булевой функцией – функция степени 2. Это определение имеет смысл, потому что, как будет предложено доказать в упражнении 9.1, любая булева функция на $\mathbb{F}_2^n$ имеет единственное полиномиальное представление в $\mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 - x_1, \ldots, x_n^2 - x_n)$.

Теорема 9.1. *Пусть M – матрица размера $bn \times bn$ над $\mathbb{F}_2$, такая, что $M = A \otimes I$, где A – матрица размера $n \times n$ над $\mathbb{F}_2$, а I – единичная матрица размера $b \times b$. Если A – ортогональная матрица, то $f : x_1 \| \ldots \| x_n \mapsto \sum_{i=1}^n q(x_i)$ – нелинейный инвариант $x \mapsto Mx$ для любой квадратичной булевой функции q.*

Доказательство. Если q – квадратичная функция, то существуют коэффициенты $c_{j,k}$, такие что $q(z) = \sum_{1\leqslant j\leqslant k\leqslant b} c_{j,k} z_j z_k$ с точностью до постоянного члена, который можно считать равным нулю. Заметим, что q может содержать линейные члены, т. к. $z_j^2 = z_j$. Обозначим $x_{i,j}$ j-й бит $x_i \in \mathbb{F}_2^b$. В точке $x = x_1 \| \ldots \| x_n$ имеем

$$f(x) = \sum_{1\leq j\leq k\leq b} c_{j,k} \sum_{i=1}^{n} x_{i,j}\, x_{i,k} = \sum_{1\leq j\leq k\leq b} c_{j,k} \begin{bmatrix} x_{1,j} & \cdots & x_{n,j} \end{bmatrix} \begin{bmatrix} x_{1,k} \\ \vdots \\ x_{n,k} \end{bmatrix}.$$

Теперь, используя тот факт, что $A^\mathsf{T}A = I$, получаем

$$f(Mx) = \sum_{1\leq j\leq k\leq b} c_{j,k} \begin{bmatrix} x_{1,j} & \cdots & x_{n,j} \end{bmatrix} A^\mathsf{T}A \begin{bmatrix} x_{1,k} \\ \vdots \\ x_{n,k} \end{bmatrix} = f(x).$$

Отсюда следует, что f является инвариантом отображения $x \mapsto Mx$. □

Теорема 9.1 приводит к нелинейному инварианту для демонстрационного шифра типа Rijndael из раздела 3.2.1.

Пример 9.4. По теореме 9.1, любая квадратичная функция вида $x_1 \| \ldots \| x_4 \mapsto \sum_{i=1}^4 q(x_i)$ является инвариантом функции `MixColumns` M шифра типа Rijndael. S-блок S обладает тем свойством, что $x \mapsto u^\mathsf{T}\mathsf{S}(x)$ квадратична при любом выборе u. Кроме того, для $u = 111$ любое $x \in \mathbb{F}_2^3$ удовлетворяет равенству

$$u^\mathsf{T}\mathsf{S}(\mathsf{S}(x)) = u^\mathsf{T}x.$$

Эти наблюдения приводят еще к одному инварианту. А именно пусть f и g – булевы функции на $\mathbb{F}_2^{96}$, определенные следующим образом:

$$f(x_1 \| \cdots \| x_{32}) = \sum_{i=1}^{32} u^\mathsf{T}x = \sum_{i=1}^{32} x_{3i-2} + x_{3i-1} + x_{3i},$$

$$g(x_1 \| \cdots \| x_{32}) = \sum_{i=1}^{32} u^\mathsf{T}\mathsf{S}(x) = \sum_{i=1}^{32} x_{3i-2}x_{3i-1} + x_{3i-2}x_{3i} + x_{3i-1}x_{3i}.$$

Раундовая функция R удовлетворяет соотношению

$$f \circ \mathsf{R} = f \circ \texttt{MixColumns} \circ \texttt{ShiftRows} \circ \texttt{SubCells} = f \circ \texttt{SubCells} = g.$$

Аналогично на R выполняется такое соотношение:

$$g \circ \mathsf{R} = g \circ \texttt{MixColumns} \circ \texttt{ShiftRows} \circ \texttt{SubCells} = g \circ \texttt{SubCells} = f.$$

Следовательно, существует множество слабых ключей, такое что $f(\mathsf{R}_{k_{i+1}}(\mathsf{R}_{k_i}(x))) = f(x)$ для любого $x \in \mathbb{F}_2^{96}$ или $f(\mathsf{R}_{k_{i+1}}(\mathsf{R}_{k_i}(x))) = f(x) + 1$ для любого x. На самом деле это справедливо для любого выбора k_{i+1}. Однако любая 3-битовая ячейка k_i должна быть равна 000 или 111, чтобы нелинейный инвариант имел место. Далее следует короткое алгебраическое преобразование полиномиального представления g (см. упражнение 9.2). Следовательно, для

2^{32} слабых ключей f является нелинейным инвариантом при любом четном числе раундов. ▷

Как было отмечено в главе 2, большое число линейных следов с малыми абсолютными корреляциями теоретически может привести к линейной аппроксимации с большой абсолютной корреляцией. Нелинейным инвариантом f из примера 9.4 фактически является линейная функция $x \mapsto u^{\mathsf{T}}x$. В частности, тот факт, что либо $f(\mathsf{E}_k(x)) = f(x)$ для всех x, либо $f(\mathsf{E}_k(x)) \neq f(x)$ для всех x эквивалентен тому, что $C^{\mathsf{E}_k}_{u,u} = \pm 1$ для $u = 11\ldots1$. В упражнении 3.2 было показано, что корреляция четырехраундовых следов не превышает 2^{-16}. Тем не менее (u, u) – линейная аппроксимация с корреляцией ± 1. Правда, это верно лишь для небольшой доли ключей ($1/2^{64}$), но может служить неплохой иллюстрацией ограничений линейных следов.

Существование нелинейных инвариантов приводит к естественному вопросу: как проанализировать корреляцию пар булевых функций, т. е. нелинейный аналог линейных аппроксимаций? В разделе 9.2.2 обсуждаются некоторые подходы к этой проблеме.

9.2. Приближенные свойства

Недостаток «точных» свойств, рассмотренных в разделе 9.1, – то, что они либо черные, либо белые; свойство либо имеет место, либо нет. Из-за этого для аппроксимаций не хватает свободы маневра.

В этом разделе мы обсудим некоторые попытки обобщить свойства из раздела 9.1, сделав их неточными. Однако все они сталкиваются со значительными трудностями, которые мы сможем рассмотреть только в главе 11. По этой причине в каждом конкретном случае мы придерживаемся описания высокого уровня.

9.2.1. Статистическое насыщение

Как было объяснено в разделе 9.1.1, атаки с насыщением основаны на шифровании множеств открытых текстов таким образом, что часть открытого текста принимает все возможные значения (эта часть называется «насыщенной»), тогда как оставшаяся часть является постоянной. Вообще, множество открытых текстов является смежным классом некоторого векторного пространства. В простейшем случае это приводит к множеству шифртекстов, таких что часть выходов обладает свойством насыщения. Однако в примере 9.1 было показано, что это также может привести к множеству шифртекстов с постоянной частью.

В статистических атаках с насыщением используется тот факт, что при шифровании случайного равномерно распределенного открытого текста из смежного класса распределение вероятностей части шифртекста неравномерно. Степень неравномерности обычно измеряется квадратичным евклидовым расхождением, поскольку эта величина определяет информационную сложность. Крайний случай, соответствующий наибольшему квадратичному евклидову расхождению, дает набор шифртекстов с постоянной частью.

Согласно следствию 6.6, свойства статистического насыщения эквивалентны многомерным линейным аппроксимациям. Разница заключается в том,

как оценивается квадратичное евклидово расхождение. На практике для этой цели обычно проще использовать линейный криптоанализ. Однако в некоторых случаях возможны более простые рассуждения, основанные на значениях.

9.2.2. Нелинейные аппроксимации

Нелинейная аппроксимация функции $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ – это пара (f, g) функций $f : \mathbb{F}_2^n \to \mathbb{F}_2$ и $g : \mathbb{F}_2^m \to \mathbb{F}_2$. Корреляция (f, g) определяется по аналогии с корреляцией линейной аппроксимации.

$$C(g \circ \mathsf{F}, f) = 2\Pr_{\mathbf{x}}[g(\mathsf{F}(\mathbf{x})) = f(\mathbf{x})] - 1,$$

где вероятность берется для случайной равномерно распределенной величины $\mathbf{x}$. Если $f = g$ и корреляция равна ± 1, то f – нелинейный инвариант F.

Использование нелинейных аппроксимаций было предложено вскоре после открытия линейного криптоанализа, но это не привело к появлению общего способа анализа нелинейных аппроксимаций. Впоследствии было предложено много других подходов; здесь мы упомянем только один из них. Если F и G – перестановки, то линейные аппроксимации $\mathsf{G} \circ \mathsf{F} \circ \mathsf{G}^{-1}$ являются нелинейными аппроксимациями F. Более того, если $\mathsf{F} = \mathsf{F}_r \circ \ldots \circ \mathsf{F}_2 \circ \mathsf{F}_1$, то «нелинейный след» для F соответствует следу для

$$\mathsf{G} \circ \mathsf{F} \circ \mathsf{G}^{-1} = (\mathsf{G} \circ \mathsf{F}_r \circ \mathsf{G}^{-1}) \circ \ldots \circ (\mathsf{G} \circ \mathsf{F}_2 \circ \mathsf{G}^{-1}) \circ (\mathsf{G} \circ \mathsf{F}_1 \circ \mathsf{G}^{-1}).$$

Излишне говорить, что это разложение не является единственным. Одним из подходов к нелинейному криптоанализу является выполнение линейного криптоанализа для такого альтернативного описания шифра.

Все вышеупомянутые предложения сталкиваются со значительными трудностями, что делает их непригодными на практике. Стоит упомянуть две повторяющиеся проблемы: (1) зависимость корреляций от ключа и (2) существует «слишком много» нелинейных функций, чтобы получить полезную теорию такого же уровня, как линейный криптоанализ. Прежде чем мы перейдем к обсуждению этих вопросов, нам придется заново выстроить теорию линейного криптоанализа в главе 11.

9.2.3. Каркас проецирования

Нелинейные аппроксимации можно обобщить на произвольные функции $f : \mathbb{F}_2^n \to X$ и $g : \mathbb{F}_2^m \to Y$, где X и Y – небольшие множества. Иногда их называют «функциями проецирования».

Общая идея криптоанализа на основе функций проецирования – соотнести $g \circ \mathsf{F}$ с f. Конкретно, в постановке с известным открытым текстом производится попытка найти сбалансированные функции[1] f и g, такие что $(f(\mathbf{x}), g(\mathsf{F}(\mathbf{x})))$ имеет неравномерное распределение для случайной равномерно распределенной величины $\mathbf{x}$. Это можно измерить с помощью квадратичного евклидова расхождения.

Альтернативная точка зрения заключается в том, что функции f и g определяют разбиения $\mathbb{F}_2^n$ и $\mathbb{F}_2^m$. Например, f разбивает $\mathbb{F}_2^n$ следующим образом:

[1] Функция называется сбалансированной, если у любого выходного значения одно и то же число прообразов.

$$\mathbb{F}_2^n = \bigcup_{x \in X} f^{-1}(x),$$

где $f^{-1}(x)$ – множество значений $y \in \mathbb{F}_2^n$, таких что $f(y) = x$. В случае *криптоанализа с разбиением* изучается связь между разбиением пространства входов и разбиением пространства выходов. Это эквивалентно криптоанализу, основанному на функциях проецирования.

Разбиение и функции проецирования позволяют описать широкий спектр свойств. Например, если f и g – линейные функции, то они эквивалентны многомерным линейным аппроксимациям. Однако существуют также некоторые заметные исключения, такие как множественные линейные аппроксимации со множеством масок, которые не образуют векторного пространства.

К сожалению, разбиение и функции проецирования способны лишь описать свойства. Они, например, никак не помогают анализировать или находить эти свойства.

9.3. Историческая справка

Свойства насыщения были введены Кнудсеном как часть «атаки Square». Их связь с многомерными линейными аппроксимациями с нулевой корреляцией заметили Богданов, Леандер, Нюберг и Вань.

Простые примеры инвариантных подпространств, такие как пример 9.2, были замечены еще до открытия инвариантных подпространств, зависящих от деталей уровня S-блоков и линейного уровня, Леандером, Абельрахимом, Аль-Хазими и Зеннером. Алгоритм 9.1 принадлежит Леандеру, Мино и Реньему. Нелинейные инварианты и теорема 9.1 введены в обиход Тодо, Леандером и Сасаки. Пример 9.4 основан на работе Бейна (2018).

Большинство приближенных свойств, рассмотренных в разделе 9.2, появились раньше точных свойств из раздела 9.1. Статистические атаки с насыщением впервые были предложены Воденэ, сам термин введен в работе Колларда и Стандаерта. Использовать нелинейные аппроксимации предложили в 1995 году Харпес, Крамер и Масси под названием «суммы ввода–вывода», а в 1996 году – Кнудсен и Робшоу. Подход, основанный на применении линейного криптоанализа к альтернативному описанию шифра, предложен Беером, Канто и Леандером. Криптоанализ с разбиением предложили Харпес и Масси, а понятие функций проецирования из раздела 9.2.3 – Ваген. По причинам, которые станут понятны в главе 11, ни одно из этих предложений не привело к жизнеспособным обобщениям криптоанализа.

9.4. Литература

Beierle, Christof, Anne Canteaut, and Gregor Leander (2018). «Nonlinear Approximations in Cryptanalysis Revisited». In: IACR *Transactions on Symmetric Cryptology* 2018.4, pp. 80–101. issn: 2519-173X. doi: 10.13154/tosc.v2018.i4.80-101.

Beyne, Tim (Dec. 2018). «Block Cipher Invariants as Eigenvectors of Correlation Matrices». In: *ASIACRYPT 2018, Part I*. Ed. by Thomas Peyrin and Steven Galbraith. Vol. 11272. LNCS. Springer, Cham, pp. 3–31. doi: 10.1007/978-3-030-03326-2_1.

Bogdanov, Andrey et al. (Dec. 2012). «Integral and Multidimensional Linear Distinguishers with Correlation Zero». In: *ASIACRYPT 2012*. Ed. by Xiaoyun Wang and Kazue Sako. Vol. 7658. LNCS. Springer, Berlin, Heidelberg, pp. 244–261. doi: 10.1007/978-3-642-34961-4_16.

Collard, Baudoin and François-Xavier Standaert (Apr. 2009). «A Statistical Saturation Attack against the Block Cipher PRESENT». In: *CT-RSA 2009*. Ed. by Marc Fischlin. Vol. 5473. LNCS. Springer, Berlin, Heidelberg, pp. 195–210. doi: 10.1007/978-3-642-00862-7_13.

Daemen, Joan, Lars R. Knudsen, and Vincent Rijmen (Jan. 1997). «The Block Cipher Square». In: *FSE'97*. Ed. by Eli Biham. Vol. 1267. LNCS. Springer, Berlin, Heidelberg, pp. 149–165. doi: 10.1007/BFb0052343.

Harpes, Carlo, Gerhard G. Kramer, and James L. Massey (May 1995). «A Generalization of Linear Cryptanalysis and the Applicability of Matsui's Piling-Up Lemma». In: *EUROCRYPT'95*. Ed. by Louis C. Guillou and Jean-Jacques Quisquater. Vol. 921. LNCS. Springer, Berlin, Heidelberg, pp. 24–38. doi: 10.1007/3-540-49264-X_3.

Harpes, Carlo and James L. Massey (Jan. 1997). «Partitioning Cryptanalysis». In: *FSE'97*. Ed. by Eli Biham. Vol. 1267. LNCS. Springer, Berlin, Heidelberg, pp. 13–27. doi: 10.1007/BFb0052331.

Knudsen, Lars R. and Matthew J. B. Robshaw (May 1996). «Non-Linear Approximations in Linear Cryptanalysis». In: *EUROCRYPT'96*. Ed. by Ueli M. Maurer. Vol. 1070. LNCS. Springer, Berlin, Heidelberg, pp. 224–236. doi: 10.1007/3-540-68339-9_20.

Leander, Gregor et al. (Aug. 2011). «A Cryptanalysis of PRINTcipher: The Invariant Subspace Attack». In: *CRYPTO 2011*. Ed. by Phillip Rogaway. Vol. 6841. LNCS. Springer, Berlin, Heidelberg, pp. 206–221. doi: 10.1007/978-3-642-22792-9_12.

Leander, Gregor, Brice Minaud, and Sondre Rшnjom (Apr. 2015). «A Generic Approach to Invariant Subspace Attacks: Cryptanalysis of Robin, iSCREAM and Zorro». In: *EUROCRYPT 2015, Part I*. Ed. by Elisabeth Oswald and Marc Fischlin. Vol. 9056. LNCS. Springer, Berlin, Heidelberg, pp. 254–283. doi: 10.1007/978-3-662-46800-5_11.

Todo, Yosuke, Gregor Leander, and Yu Sasaki (Dec. 2016). «Nonlinear Invariant Attack – Practical Attack on Full SCREAM, iSCREAM, and Midori64». In: *ASIACRYPT 2016, Part II*. Ed. by Jung Hee Cheon and Tsuyoshi Takagi. Vol. 10032. LNCS. Springer, Berlin, Heidelberg, pp. 3–33. doi: 10.1007/978-3-662-53890-6_1.

Vaudenay, Serge (Mar. 1996b). «An Experiment on DES Statistical Cryptanalysis». In: *ACM CCS 96*. Ed. by Li Gong and Jacques Stern. ACM Press, New York, pp. 139–147. doi: 10.1145/238168.238206.

9.5. Упражнения

Упражнение 9.1

Цель этого упражнения – показать, что любой функции $f : \mathbb{F}_2^n \to \mathbb{F}_2$ соответствует единственный полином, принадлежащий $\mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 - x_1, \ldots, x_n^2 - x_n)$. Этот полином называется *алгебраической нормальной формой f*.

1. Покажите, что любой полином, принадлежащий $\mathbb{F}_2[x_1, \ldots, x_n]$, определяет булеву функцию путем вычисления.

2. Покажите, что для любой булевой функции существует интерполирующий полином в $\mathbb{F}_2[x_1, \ldots, x_n]$.
3. Применив рассуждение с подсчетом, сделайте вывод, что «отображение вычисления», которое переводит полином в булеву функцию, определяемую вычислением полинома, является взаимно однозначным отображением между $\mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 - x_1, \ldots, x_n^2 - x_n)$ и множеством всех булевых функций.

Упражнение 9.2

Для каких ключей нелинейный инвариант из примера 9.4 имеет место при произвольном четном числе раундов? Докажите.

Упражнение 9.3

Знаменитый бельгийский криптограф часто шифрует свои персональные данные своим любимым блочным шифром. У этого блочного шифра есть три варианта с $r_1 = 10$, $r_2 = 12$ и $r_3 = 14$ раундами. К сожалению, в данном случае криптограф не помнит, какой вариант использовал.

Но, к счастью, криптограф попросил своих студентов записывать для него число раундов. Однако в творческом порыве студенты решили зашифровать это число придуманным ими шифром E_k с размером блока 4 бита. Как показано на рис. 9.2, на i-м раунде их построения к состоянию прибавляется i-й фрагмент k_i ключа $k = k_1 \| k_2 \| \ldots \| k_{r+1}$, а затем применяется функция S, заданная табл. 9.1. Не слишком доверяя собственным способностям, студенты решили создать экземпляр своего шифра E_k с $r = r_1 \times r_2 \times r_3 + 1 = 1681$ раундом.

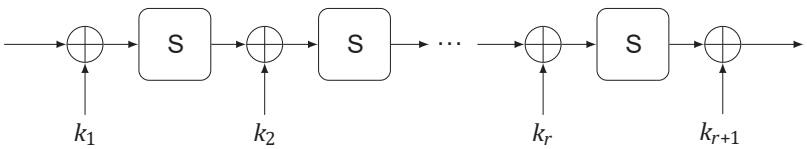


Рис. 9.2. Студенческий метод шифрования

Таблица 9.1. Справочная таблица для функции S

0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
3	e	6	8	0	c	b	4	1	d	5	a	7	9	f	2

Студенты записали, что результат шифрования $r_1 = 10$ равен $\mathsf{E}_k(1,0,1,0) = (0,1,0,1)$, т. е. пяти. Они также помнят, что $r_2 = 12$ соответствует шифртекст $\mathsf{E}_k(1,1,0,0) = (0,0,0,0)$. Разумеется, студенты забыли ключ, но помнят, что это был ASCII-код парольной фразы, состоящей только из строчных и заглавных букв. Услышав это, знаменитый криптограф воскликнул, что студенты допустили ошибку. Можете ли вы помочь студентам выяснить, в чем они неправы?

Глава 10

Функции на абелевых группах

В главе 11 теория линейного криптоанализа пересматривается с более общей точки зрения. Для этого нам понадобятся некоторые знания из математики. Сначала мы обсудим линейную алгебру над полем комплексных чисел, а затем обратимся к анализу Фурье на конечной абелевой группе. Обе эти темы играют центральную роль в главе 11.

10.1. Линейная алгебра над полем $\mathbb{C}$

Линейная алгебра занимается векторными пространствами и линейными преобразованиями между ними. Однако если векторное пространство определено над полем вещественных или комплексных чисел, то появляется дополнительная структура, связанная с тем, что в поле $\mathbb{C}$ определена функция абсолютной величины, или модуля, которая превращает его в метрическое пространство.

В этой главе, как и в главе 11, важную роль играют два векторных пространства над $\mathbb{C}$. Они используются далее в качестве сквозного примера.

Пример 10.1. Пусть G – конечная абелева группа. Свободное $\mathbb{C}$-векторное пространство на G состоит из множества всех формальных линейных комбинаций элементов G. То есть каждый элемент $\mathbb{C}[G]$ имеет вид

$$\sum_{x \in G} u_x \, \delta_x,$$

где значениями u_x являются произвольные комплексные числа, а δ_x – формальный базисный вектор, соответствующий элементу группы x. Базис $\{\delta_x \mid x \in G\}$ называется стандартным базисом $\mathbb{C}[G]$.

Аналогично векторное пространство $\mathbb{C}^G$ состоит из всех функций $G \to \mathbb{C}$. Если δ_x обозначает функцию, равную 1 в точке x и 0 во всех остальных точках, то любую функцию $f \in \mathbb{C}^G$ можно записать в виде

$$\sum_{x \in G} f(x) \, \delta^x.$$

В упражнении 10.1 вам будет предложено проверить, что $\{\delta_x \mid x \in G\}$ является базисом $\mathbb{C}^G$. Он называется стандартным базисом $\mathbb{C}^G$. Оба векторных пространства $\mathbb{C}[G]$ и $\mathbb{C}^G$ изоморфны $\mathbb{C}^{|G|}$. ▷

10.1.1. Нормированные векторные пространства и двойственные им

Векторное пространство над $\mathbb{C}$ можно снабдить нормой, являющейся абстракцией понятия «длины».

Определение 10.1 (нормированное векторное пространство). Пусть V – векторное пространство над $\mathbb{C}$. Нормой на V называется вещественная функция $\|\cdot\|\colon V \to \mathbb{R}$ на V такая, что

(1) для любого $x \in V$ $\|x\| \geqslant 0$, причем равенство имеет место тогда и только тогда, когда $x = 0$;
(2) для любых $x \in V$ и $\lambda \in \mathbb{C}$ $\|\lambda x\| = |\lambda|\,\|x\|$;
(3) выполняется неравенство треугольника: для любых $x, y \in V$ $\|x + y\| \leqslant \|x\| + \|y\|$.

Векторное пространство с нормой называется нормированным векторным пространством.

Пример 10.2. Для любого $p \in [1, \infty)$ векторное пространство $\mathbb{C}[G]$ можно снабдить так называемой p-нормой $\|\cdot\|_p$. p-норма вектора u с координатами u_x для $x \in G$ определяется формулой

$$\|u\|_p = \sqrt[p]{|G|}\left(\sum_{x \in G} |u_x|^p\right)^{\frac{1}{p}}.$$

Похожая норма, которая, допуская некоторую вольность нотации, также обозначается $\|\cdot\|_p$, определяется на $\mathbb{C}^G$ формулой

$$\|f\|_p = \frac{1}{\sqrt[p]{|G|}}\left(\sum_{x \in G} |f(x)|^p\right)^{\frac{1}{p}}.$$

Для $p = 2$ это хорошо знакомая *евклидова норма*. Проверьте, что она действительно является нормой. Для иллюстрации некоторых идей полезен общий случай, но в главе 11 используется только евклидова норма. Поэтому мы опускаем доказательство того, что p-норма действительно удовлетворяет свойствам, указанным в определении 10.1, для всех $p \geqslant 1$. ▷

Для любого векторного пространства имеется двойственное ему векторное пространство. При этом векторное пространство, двойственное нормированному, само является нормированным. В следующем определении предполагается, что V конечномерно, чтобы избежать топологических тонкостей.

Определение 10.2 (двойственное векторное пространство). Пусть V – конечномерное векторное пространство над $\mathbb{C}$ с нормой $\|\cdot\|$. Двойственным пространством $V^\vee$ пространства V называется $\mathbb{C}$-векторное пространство всех линейных функций $V \to \mathbb{C}$ с нормой

$$\|f\|^\vee = \max_{\substack{v \in V \\ \|v\| \leq 1}} |f(v)|.$$

Элементы $V^\vee$ называются линейными функционалами.

Проверьте, что двойственная норма из определения 10.2 действительно является нормой. Так как $\dim V$ и $\dim V^\vee$ равны, векторные пространства V и $V^\vee$ изоморфны. Действительно, базис V можно отобразить в базис $V^\vee$. Выбор изоморфизма произволен, потому что разные базисы обычно приводят к разным изоморфизмам. Кроме того, такие изоморфизмы, вообще говоря, не являются изометриями, т. е. не сохраняют норму. Однако существует «канонический» изометрический изоморфизм между V и $V^{\vee\vee}$, который можно определить, не прибегая к такому произвольному выбору базиса.

Теорема 10.3. *Пусть V – конечномерное векторное пространство над $\mathbb{C}$ с нормой $\|\cdot\|$. Для любого $v \in V$ определим «отображение вычисления»* $\mathrm{ev}_v : V^\vee \to \mathbb{C}$ как $\mathrm{ev}_v(f) \quad f(v)$. *Функция $V \to V^{\vee\vee} : v \mapsto \mathrm{ev}_v$ является изоморфизмом векторных пространств. Более того, она является изометрией нормированных векторных пространств:* $\|\mathrm{ev}_v\|^{\vee\vee} = \|v\|$.

Доказательство. Нетрудно видеть, что $\mathrm{ev}_{\lambda v} = \lambda\,\mathrm{ev}_v$ и $\mathrm{ev}_{u+v} = \mathrm{ev}_u + \mathrm{ev}_v$. Поэтому $v \mapsto \mathrm{ev}_v$ является гомоморфизмом векторных пространств. Его ядро нулевое, а сравнение размерностей показывает, что это должен быть изоморфизм векторных пространств. Чтобы показать, что он является изометрией, сначала докажем, что верхняя граница $\|\mathrm{ev}_v\|^{\vee\vee}$ равна:

$$\|\mathrm{ev}_v\|^{\vee\vee} = \max_{\substack{f \in V^\vee \\ \|f\|^\vee \le 1}} |f(v)| \le \|v\|,$$

где мы воспользовались тем фактом, что $|f(v)| \leqslant \|f\|^\vee \|v\|$. Это следует из определения $\|f\|^\vee$. Кроме того, обязательно существует функционал f с $\|f\|^\vee \leqslant 1$, такой что $|f(v)| = \|v\|$. Действительно, положим $f(\alpha v) = \alpha\|v\|$ на $\mathrm{Span}\{v\}$ и продолжим его на все V. Такое продолжение всегда возможно, потому что V конечномерно, поэтому V имеет базис, содержащий v. □

Пример 10.3. Векторное пространство, двойственное $\mathbb{C}[G]$, состоит из всех линейных функций из $\mathbb{C}[G]$ в $\mathbb{C}$. Однако любая линейная функция $f : \mathbb{C}[G] \to \mathbb{C}$ определяется своим образом на базисных векторах δ_x, где $x \in G$:

$$f\left(\sum_{x\in G} u_x\,\delta_x\right) = \sum_{x\in G} u_x\,f(\delta_x).$$

Следовательно, линейные функции на $\mathbb{C}[G]$ эквивалентны элементам $\mathbb{C}^G$. С точностью до этого канонического изоморфизма $\mathbb{C}_G$ совпадает с $\mathbb{C}[G]^\vee$. Если $\mathbb{C}[G]$ снабжено p-нормой, то двойственная норма на $\mathbb{C}_G$ является $p/(p-1)$-нормой. В упражнении 10.2 вам будет предложено доказать этот факт. Случай $p = 2$ особый: норма, двойственная евклидовой, сама является евклидовой нормой. Отсюда следует, что отображение

$$f \mapsto \frac{1}{|G|}\sum_{x\in G} f(x)\,\delta_x$$

является изометрическим изоморфизмом между $(\mathbb{C}^G, \|\cdot\|_2)$ и $(\mathbb{C}[C], \|\cdot\|_2)$. Так как нормированное векторное пространство $(\mathbb{C}[G], \|\cdot\|_2)$ изометрически изоморфно своему двойственному, оно называется самодвойственным. ▷

10.1.2. Пространства со скалярным произведением

В примере 10.3 показано, что евклидова норма является самодвойственной. Это также следует из того, что она индуцирована скалярным произведением.

Определение 10.4 (пространство со скалярным произведением). Пусть V – векторное пространство над $\mathbb{C}$. Скалярным произведением на V называется функция $V \times V \to \mathbb{C}$, обозначаемая $\langle\cdot,\cdot\rangle$, такая что:

(1) для любых $x, y, z \in V$ и $\lambda, \mu \in \mathbb{C}$ $\langle x, \lambda y + \mu z\rangle = \lambda\langle x, y\rangle + \mu\langle x, z\rangle$;
(2) она антисимметрична: $\overline{\langle x, y\rangle} = \langle y, x\rangle$ для любых $x, y \in V$;
(3) для любого $x \in V$ $\langle x, x\rangle \geqslant 0$, причем равенство имеет место тогда и только тогда, когда $x = 0$.

Следующий результат показывает, что любое пространство со скалярным произведением является нормированным.

Теорема 10.5. *Если V – векторное пространство со скалярным произведением $\langle\cdot,\cdot\rangle$, то $x \mapsto \|x\| = \sqrt{\langle x, x\rangle}$ является нормой на V.*

Более того, всякое пространство со скалярным произведением является самодвойственным в смысле примера 10.3. В следующей теореме антиизоморфизмом векторных пространств над $\mathbb{C}$ называется обратимое отображение f, такое что $f(\lambda x + \mu y) = \bar{\lambda} f(x) + \bar{\mu} f(y)$ для всех x и y и скаляров λ и μ. Необходимость в антиизоморфизмах объясняется требованиями антисимметричности и положительной определенности скалярного произведения.

Теорема 10.6. *Пусть V – конечномерное векторное пространство со скалярным произведением $\langle\cdot,\cdot\rangle$. Для любого $x \in V$ определим $x^* \in V^\vee$ как $x^*(y) = \langle x, y\rangle$ для всех $y \in V$. Отображение $x \mapsto x^*$ является изометрическим антиизоморфизмом.*

Доказательство. См. упражнение 10.4. □

Пример 10.4. Стандартное скалярное произведение на $\mathbb{C}_G$ определяется следующим образом:

$$\langle f, g\rangle = \frac{1}{|G|} \sum_{x \in G} \overline{f(x)}\, g(x).$$

Для $G = \mathbb{F}_2^n$ это скалярное произведение, которое мы использовали в упражнении 2.1. Похожее скалярное произведение определено на $\mathbb{C}[G]$:

$$\langle u, v\rangle = |G| \sum_{x \in G} \overline{u}_x\, v_x,$$

где u и v – векторы с координатами u_x и v_x, $x \in G$, соответственно.

В силу теорем 10.3 и 10.6, эти скалярные произведения индуцируют изометрические антиизоморфизмы между $\mathbb{C}^G$ и $\mathbb{C}[G]$. Следовательно, если рассматривать только структуру евклидовой нормы, то эти пространства во всех отношениях неразличимы. ▷

Пространства со скалярным произведением допускают геометрическую интерпретацию. Говорят, что два вектора ортогональны, если их скалярное произведение равно нулю. То есть $u \perp v$ тогда и только тогда, когда $\langle u, v\rangle = 0$. Базис, состоящий из взаимно ортогональных векторов с единичной нормой, называется ортонормированным.

Вообще, модель скалярного произведения двух нормированных векторов можно интерпретировать как косинус наименьшего угла между ними – хотя для векторов, отличных от вещественных, некоторые предпочитают определять угол как вещественную часть скалярного произведения. В упражнении 10.7 понятие угла между векторами обобщается на угол между двумя подпространствами.

Теорема 10.7 (*теорема Пифагора). Для любой пары ортогональных векторов u и v в пространстве со скалярным произведением* $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

Доказательство. Результат следует из того, что $\|u + v\|^2 = \langle u + v, u + v\rangle$ и $\langle u + v, u + v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle = \|u\|^2 + \|v\|^2 + \langle u, v\rangle + \overline{\langle u, v\rangle}$.

Так как u и v ортогональны, $\langle u, v\rangle = 0$. □

Ортогональным дополнением подпространства V пространства со скалярным произведением U называется векторное пространство $V^\perp$, состоящее из всех векторов, ортогональных V:

$$V^\perp = \{u \in U \mid \langle v, u\rangle = 0 \text{ для любого } v \in V\}.$$

В упражнении 10.5 вам нужно будет показать, что ортогональные дополнения являются также алгебраическими дополнениями. То есть $U = V \oplus V^\perp$, где $\oplus$ обозначает внутреннюю прямую сумму. Следовательно, можно определить проекцию $\pi_V : U \to V$ с ядром $V^\perp$. Для любого $u \in U$ $\pi_V(u)$ называется *ортогональной проекцией* u на V.

Следующий результат, который иногда называют *теоремой о наилучшей аппроксимации*, важен в связи с пространствами со скалярным произведением.

Теорема 10.8. *Пусть V – подпространство конечномерного пространства со скалярным произведением U. Ортогональная проекция $u \in U$ на V является точкой, ближайшей к u:*

$$\|u - \pi_V(u)\| = \min_{v \in V} \|u - v\|.$$

Кроме того, если u не ортогонален V, то

$$\frac{|\langle \pi_V(u), u\rangle|}{\|\pi_V(u)\|} = \|\pi_V(u)\| = \max_{\substack{v \in V \\ v \neq 0}} \frac{|\langle v, u\rangle|}{\|v\|}.$$

Доказательство. Первое утверждение состоит в том, что ортогональная проекция минимизирует норму. Для любого $v \in V$ из теоремы Пифагора следует, что

$$\|u - v\|^2 = \|u - \pi_V(u) + v - \pi_V(u)\|^2 = \|u - \pi_V(u)\|^2 + \|v - \pi_V(u)\|^2.$$

Следовательно, если $v \neq \pi_V(u)$, то $\|u - v\| > \|u - \pi_V(u)\|$. Вторая часть теоремы вытекает из того, что любой ненулевой вектор $v \in V$ удовлетворяет неравенству

$$\frac{|\langle u, v\rangle|}{\|v\|} = \frac{|\langle \pi_V(u), v\rangle|}{\|v\|} \leq \|\pi_V(u)\|.$$

На этом доказательство завершается. □

10.1.3. Сингулярное разложение

Сопряженным линейному отображению $L : U \to V$ между пространствами со скалярным произведением U и V называется линейное отображение $L^\dagger : V \to U$, однозначно определяемое соотношением

$$\langle L^\dagger(v), u\rangle = \langle v, L(u)\rangle$$

для любых $u \in U$ и $v \in V$. Матричным представлением $L^\dagger$ относительно двух базисов будет сопряженно-транспонированная матрица L относительно тех же базисов. Это следует из антисимметричности скалярного произведения.

Линейное отображение $L^\dagger L$ является самосопряженным: $(L^\dagger L)^\dagger = L^\dagger L$. Важный результат линейной алгебры заключается в том, что самосопряженные отображения диагонализуемы относительно ортогонального базиса (см. упражнение 10.8). То есть существует ортонормированный базис $u_1, \ldots, u_d$, состоящий из собственных векторов $L^\dagger L$. Поскольку $\langle L(u), L(u)\rangle \geqslant 0$, соответствующие собственные значения являются неотрицательными вещественными числами $\sigma_1^2 \geqslant \ldots \geqslant \sigma_d^2$. Положим $v_i = L(u_i)/\sigma_i$ для $\sigma_i \neq 0$ и дополним до ортонормированного базиса $v_1, \ldots, v_d$ образа L. Это показывает, что любое линейное отображение $L : U \to V$ имеет *сингулярное разложение*.

Определение 10.9 (сингулярное разложение). Пусть $L : U \to V$ – линейное отображение между пространствами со скалярным произведением U и V, и пусть $\sigma_1^2 \geqslant \ldots \geqslant \sigma_d^2$ – собственные значения $L^\dagger L$. Сингулярное разложение L состоит из ортонормированных базисов $\{u_1, \ldots, u_d\}$ и $\{v_1, \ldots, v_d\}$ пространств U и V соответственно таких, что

$$L(x) = \sum_{i=1}^{d} \sigma_i \langle u_i, x\rangle v_i$$

для любого $x \in U$. Векторы $u_1, \ldots, u_d$ и $v_1, \ldots, u_d$ называются левым и правым сингулярными векторами соответственно.

10.1.4. Тензорные произведения векторных пространств

Тензорным произведением $\mathbb{C}$-векторных пространств U и V называется $\mathbb{C}$-векторное пространство $U \otimes V$ в совокупности с билинейным отображением $\otimes$: $U \times V \to U \otimes V$, обладающим «универсальным свойством», – оно единственным образом линеаризует произвольные билинейные отображения. Точнее, для любого отображения $T : U \times V \to W$, линейного по каждой переменной (билинейного), существует единственное *линейное* отображение $L : U \otimes V \to W$ такое, что $T(u, v) = L(u \otimes v)$.

Это еще не определяет однозначно тензорное произведение двух векторных пространств. Однако если два векторных пространства $U \otimes_1 V$ и $U \otimes_2 V$ обладают вышеупомянутым универсальным свойством, то существует единственный изоморфизм $\theta : U \otimes_1 V \to U \otimes_2 V$, такой что $\otimes_2 = \theta \circ \otimes_1$. Именно поэтому мы можем говорить о тензорном произведении, не опасаясь двусмысленности.

Это стандартное определение тензорного произведения, но оно абстрактное. В конкретных случаях удобно работать с конкретным построением тензорного произведения. Следующий пример иллюстрирует эту мысль для $\mathbb{C}[G]$ и $\mathbb{C}^G$.

Пример 10.5. Свободное векторное пространство $\mathbb{C}[G^2]$ на парах элементов G является тензорным произведением $\mathbb{C}[G]$ с собой же, где отображение $\otimes$ определено как $\delta_x \otimes \delta_y = \delta_{(x,y)}$. В явном виде $\mathbb{C}[G] \otimes \mathbb{C}[G] = \mathbb{C}[G^2]$.

Аналогично векторное пространство $\mathbb{C}^{G^2}$ функций двух переменных на G является тензорным произведением $\mathbb{C}_G$ с собой, и в этом случае $\otimes$ определено как

$$(f \otimes g)(x, y) = f(x)\, g(y).$$

В разделе 10.2 и главе 11 под *тензорным произведением* всегда будет пониматься одно из этих двух конкретных построений. ▷

Элементы тензорного произведения двух векторных пространств иногда называют тензорами, а элементы вида $u \otimes v$ – *элементарными* тензорами, или тензорами *первого ранга*.

Поскольку линейные функции из одного векторного пространства U в другое векторное пространство V сами образуют векторное пространство, тензорное произведение линейных отображений корректно определено. Тензорное произведение $L_1 \otimes \ldots \otimes L_n$ линейных отображений $L_i : V_i \to U_i$ можно канонически идентифицировать с помощью линейного отображения

$$\bigotimes_{i=1}^{n} V_i \to \bigotimes_{i=1}^{n} U_i$$
$$v_1 \otimes \cdots \otimes v_n \mapsto (L_1 v_1) \otimes \cdots \otimes (L_n v_n).$$

Матричное представление $L_1 \otimes \ldots \otimes L_n$ относительно базисов $\otimes_{i=1}^{n} U_i$ и $\otimes_{i=1}^{n} V_i$, состоящее из тензоров первого ранга, является произведением Кронекера (см. главу 2) n матриц.

10.2. Анализ Фурье на конечных абелевых группах

Этот раздел посвящен пространству со скалярным произведением $\mathbb{C}^G$ функций на конечной абелевой группе G. Если заданы функция $f \in \mathbb{C}^G$ и константа $t \in G$, то можно определить новую функцию $x \mapsto f(x + t)$ с помощью сдвига. По-другому можно сказать, что группа G действует на $\mathbb{C}^G$. Как показано в этом разделе, это действие естественно приводит к преобразованию Фурье.

10.2.1. Характеры группы

Воздействие переносов на координаты функций в стандартном базисе $\mathbb{C}_G$ неудобно: базисные векторы меняются местами перестановкой $\delta^x \mapsto \delta^{x-t}$, что соответствует умножению на матрицу перестановки. Было бы удобнее, если бы воздействие сдвига сводилось к простому масштабированию координат, т. е. умножению на диагональную матрицу. Этого можно достичь, если работать в другом базисе.

Чтобы диагонализировать действие группы, новые базисные векторы должны быть собственными векторами множества операторов сдвига. Заранее не ясно, разделяют ли эти операторы общий базис собственных векторов. Оказывается, что это так, только если группа G абелева. Функция $\chi : G \to \mathbb{C}$ является общим собственным вектором для всех сдвигов тогда и только тогда, когда $\chi(x + t) = \chi(t)\chi(x)$ для любых $x, t \in G$. Иными словами, χ должна быть гомоморфизмом из G в мультипликативную группу $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$. Это ведет к следующему определению.

Определение 10.10 (двойственная группа). Пусть G – конечная абелева группа. Комплексным характером G называется гомоморфизм групп $G \to \mathbb{C}^\times$. Двойственной в смысле Понтрягина к группе G называется группа $\hat{G}$ всех характеров G с операцией поточечного произведения.

Поточечным произведением двух характеров χ и ψ называется характер $x \mapsto \chi(x)\psi(x)$. В каждой группе имеется тривиальный характер $x \mapsto 1$, который действует как нейтральный элемент поточечного умножения. Обратным к χ является характер $x \mapsto \chi(-x) = \chi(x)$. Поэтому G действительно является группой, как и заявлено в определении 10.10.

Происхождение терминологии «двойственная группа» объясняется следующей теоремой, аналогичной теореме 10.3.

Теорема 10.11. *Пусть G – конечная абелева группа. Для любого $x \in G$ определим «отображение вычисления» $ev_x : G \to \mathbb{C}$ как $ev_x(\chi) = \chi(x)$. Любое отображение вычисления ev_x определяет характер G. Кроме того, функция $x \mapsto ev_x$ является изоморфизмом групп, отображающим G в группу, двойственную двойственной G.*

Доказательство. См. упражнение 10.9. □

Оказывается, что G и $\hat{G}$ изоморфны с той оговоркой, что канонического выбора изоморфизма не существует. В примере 10.6 вычисляется группа, двойственная циклической. Уже здесь устанавливается частный случай результата.

Пример 10.6 (группа, двойственная циклической). Обозначим $\mathbb{Z}_n$ аддитивную группу целых чисел по модулю n. Так как $nx = 0$ для любого $x \in \mathbb{Z}_n$, для каждого характера χ имеет место равенство $\chi(x)^n = 1$. Поэтому $\hat{\mathbb{Z}}_n$ является группой с экспонентой, не превышающей n. Кроме того, $\chi(x) = \chi(1)^x$ для любого $x \in \mathbb{Z}_n$. Отсюда следует, что $\hat{\mathbb{Z}}_n$ – циклическая группа порядка, не превышающего n.

Обозначим ζ примитивный корень n-й степени из единицы, такой что $\zeta = e^{2\pi\sqrt{-1}/n}$. Всякая функция $\chi_u : x \mapsto \zeta^{ux}$ является характером $\mathbb{Z}_n$, потому что $\chi_u(0) = 1$ и

$$\chi_u(x + y) = \zeta^{u(x+y)} = \zeta^{ux}\zeta^{uy} = \chi_u(x)\,\chi_u(y).$$

Выше было показано, что порядок $\hat{\mathbb{Z}}_n$ не превышает n, поэтому функции $x \mapsto \zeta^{ux}$ являются единственными характерами $\hat{\mathbb{Z}}_n$. На самом деле $u \mapsto \chi_u$ – изоморфизм между $\mathbb{Z}_n$ и двойственной ей группой. Этот изоморфизм зависит от выбора ζ. ▷

Прямой суммой $G \oplus H$ групп G и H называется группа с множеством элементов $G \times H$ и операцией $(a, b) + (c, d) = (a + c, b + d)$. Следующая теорема описывает структуру $\widehat{G \oplus H}$.

Теорема 10.12. *Пусть G и H – конечные абелевы группы. Существует изоморфизм между $\widehat{G \oplus H}$ и $\hat{G} \oplus \hat{H}$:*

$$\chi \mapsto (\chi_G, \chi_H),$$

где χ_G и χ_H обозначают ограничение χ на G и на H соответственно.

Доказательство. Обозначим f отображение, определенное в теореме. Это действительно гомоморфизм, потому что $f(\chi\psi) = ((\chi\psi)_G, (\chi\psi)_H) = (\chi_G\psi_G, \chi_H\psi_H) = (\chi_G, \chi_H)(\psi_G, \psi_H)$. Кроме того, это биекция, потому что обратное отображение имеет вид

$$g : (\chi, \psi) \mapsto (x, y) \mapsto \chi(x)\psi(y).$$

Действительно, $g(\chi, \psi)_G = \chi$ и $g(\chi, \psi)_H = \psi$. □

Основная теорема конечных абелевых групп утверждает, что любая конечная абелева группа изоморфна прямой сумме циклических групп. То есть

$$G \cong \bigoplus_n \mathbb{Z}_n.$$

В силу теоремы 10.12 и примера 10.6, отсюда следует, что

$$\widehat{G} \cong \bigoplus_n \widehat{\mathbb{Z}}_n \cong \bigoplus_n \mathbb{Z}_n \cong G.$$

Следовательно, всякая конечная абелева группа изоморфна двойственной себе. Важнее, впрочем, то, что путем выбора специального изоморфизма и следуя по приведенной выше цепочке рассуждений в обратном направлении, элементы двойственной группы можно найти явно.

Пример 10.7. Группа $\mathbb{F}_2^n$ изоморфна $\mathbb{F}_2 \oplus \mathbb{F}_2 \oplus \ldots \oplus \mathbb{F}_2$. Один из возможных изоморфизмов – отображение $x \mapsto (x_1, x_2, \ldots, x_n)$, где x_i – i-я координата x в стандартном базисе. В силу примера 10.6, $\mathbb{F}_2 = \{\psi_0, \psi_1\}$, где $\psi_u(x) = (-1)^{ux}$. Следовательно, обращение изоморфизма из теоремы 10.12 показывает, что характерами $\mathbb{F}_2^n$ являются

$$\chi_u(x) = \prod_{i=1}^{n} \psi_{u_i}(x) = (-1)^{u^{\mathsf{T}} x}.$$

В частности, $u \mapsto \chi_u$ является изоморфизмом между группой $\mathbb{F}_2^n$ и двойственной ей. ▷

Напомним, что первоначальной причиной для введения двойственной группы было то, что характеры группы являются собственными векторами

действия сдвига G на $\mathbb{C}_G$. Число характеров G равно $|G|$. Кроме того, из следующей теоремы вытекает линейная независимость характеров. Стало быть, характеры образуют полный базис $\mathbb{C}_G$, состоящий из собственных векторов.

Теорема 10.13 (ортогональность характеров). *Для любых характеров χ и ψ конечной абелевой группы G имеет место следующее:*

$$\langle \chi, \psi \rangle = \begin{cases} 1, & \text{если } \chi = \psi, \\ 0 & \text{в противном случае.} \end{cases}$$

Иными словами, характеры образуют ортонормированный базис $\mathbb{C}_G$.

Доказательство. По определению скалярного произведения на $\mathbb{C}_G$,

$$\langle \chi, \psi \rangle = \frac{1}{|G|} \sum_{x \in G} \overline{\chi(x)}\, \psi(x) = \frac{1}{|G|} \sum_{x \in G} (\psi/\chi)(x),$$

где второе равенство следует из того факта, что $\bar{\chi}$ является обращением χ. Если $\psi = \chi$, то $\psi/\chi \equiv 1$ и скалярное произведение равно 1. Если $\psi \neq \chi$, то существует значение $t \in G$, такое что $(\psi/\chi)(t) \neq 1$. Отсюда

$$\frac{1}{|G|} \sum_{x \in G} (\psi/\chi)(x) = \frac{1}{|G|} \sum_{x \in G} (\psi/\chi)(x + t) = \underbrace{(\psi/\chi)(t)}_{\neq 1} \frac{1}{|G|} \sum_{x \in G} (\psi/\chi)(x).$$

Из этого равенства следует, что скалярное произведение равно нулю. □

10.2.2. Преобразование Фурье

Преобразование Фурье – это, по существу, переход от базиса характеров к стандартному базису. Однако чтобы избежать выбора произвольного изоморфизма между $\hat{G}$ и G, лучше определить его как преобразование $\mathbb{C}^G$ в $\mathbb{C}^{\hat{G}}$. При таком определении преобразование Фурье отображает характер $\chi \in \hat{G} \subset \mathbb{C}^G$ непосредственно в вектор стандартного базиса $\delta^\chi \in \mathbb{C}^{\hat{G}}$. Поскольку характеры групп ортогональны (по теореме 10.13), определение 10.14 дает желаемое преобразование базиса.

Определение 10.14 (преобразование Фурье). Пусть $f\colon G \to \mathbb{C}$ – некоторая функция. Преобразованием Фурье f называется функция $\hat{f}\colon \hat{G} \to \mathbb{C}$, определенная следующим образом:

$$\hat{f}(\chi) = \langle \chi, f \rangle = \frac{1}{|G|} \sum_{x \in G} \overline{\chi(x)} f(x).$$

Преобразование Фурье – это отображение $\mathcal{F} : \mathbb{C}^G \to \mathbb{C}^{\hat{G}}$, определенное как $\mathcal{F}(f) = \hat{f}$.

Как было сказано выше, преобразование Фурье $\mathcal{F}$ из определения 10.14 отображает каждый характер χ в соответствующую функцию стандартного базиса δ^χ. Напомним (см. пример 10.3), что $\mathbb{C}^G$ двойственно $\mathbb{C}[G]$. В частности, для базиса характеров существует *двойственный базис*, состоящий из векторов χ^*, определенных как

$$\chi^* = \frac{1}{|G|} \sum_{x \in G} \overline{\chi(x)}\, \delta_x.$$

Заметим, что χ^* – результат применения антиизоморфизма из теоремы 10.6 к χ. По теореме 10.13, $\psi(\chi^*) = \langle \chi, \psi \rangle = 1$, если $\chi = \psi$, и нулю в противном случае. Это свойство, к которому апеллирует термин «двойственный базис». Соответственно, существует двойственная версия преобразования Фурье, определенная путем отображения каждого вектора χ^* в соответствующий вектор стандартного базиса δ^χ. В силу ортогональности характеров (теорема 10.13), следующее определение реализует это преобразование.

Определение 10.15 (преобразование Фурье, двойственное). Пусть u – вектор, принадлежащий $\mathbb{C}[G]$. Преобразованием Фурье u называется вектор $\hat{u} \in \mathbb{C}[\hat{G}]$, определенный как

$$\widehat{u}_\chi = \sum_{x \in G} \chi(x)\, u_x.$$

Преобразованием Фурье называется отображение $\mathrm{F}\colon \mathbb{C}[G] \to \mathbb{C}[\hat{G}]$, определенное как $\mathrm{F}(u) = \hat{u}$.

«Двойственное» преобразование Фурье F из определения 10.15 связано с преобразованием Фурье $\mathcal{F}$ из определения 10.14 соотношением $\mathrm{F} = \mathcal{F}^{-\mathrm{v}}$. Здесь $\mathcal{F}^{\mathrm{v}}\colon \mathbb{C}[\hat{G}] \to \mathbb{C}[G]$ – результат транспонирования $\mathcal{F}$, определенного соотношением

$$f(\mathcal{F}^{\mathrm{v}}(u) = \mathcal{F}(f)(u)$$

для любых $u \in \mathbb{C}[\hat{G}]$ и $f \in \mathbb{C}^G$. Если положить в этом равенстве $u = \mathcal{F}^{-\mathrm{v}}(\chi^*)$ и $f = \psi$ равной характеру G, то получим $\delta_\psi(u) = \psi(\chi^*)$. Отсюда $u = \delta^\chi$ и $\mathrm{F} = \mathcal{F}^{-\mathrm{v}}$.

Пример 10.8. Преобразование Фурье из определения 5.2 является преобразованием Фурье на $\mathbb{C}[\mathbb{F}_2^n]$. Действительно, в силу примера 10.7, характерами $\mathbb{F}_2^n$ являются

$$\chi_u(x) = (-1)^{u^{\mathrm{T}}x}.$$

Подстановка этого равенства в определение 10.15 дает определение 5.2. ▷

Векторное пространство $\mathbb{C}^{\hat{G}}$ является пространством со скалярным произведением

$$\langle f, g \rangle = \sum_{\chi \in \widehat{G}} \overline{f(\chi)} g(\chi)$$

для любых $f, g \in \mathbb{C}^{\hat{G}}$. При таком выборе скалярного произведения преобразование Фурье унитарно. Аналогичный результат имеет место для преобразования F.

Теорема 10.16. *Преобразование Фурье $\mathcal{F} : \mathbb{C}^{\mathrm{G}} \to \mathbb{C}^{\hat{G}}$ унитарно. Это означает, что $\mathcal{F}^{-1} = \mathcal{F}^\dagger$, где $\mathcal{F}^\dagger$ – преобразование, сопряженное $\mathcal{F}$. В явном виде если $\hat{f} = \mathcal{F}(f)$, то*

$$f(x) = \sum_{\chi \in \widehat{G}} \chi(x)\, \widehat{f}(\chi).$$

Доказательство. По определению, преобразование, сопряженное $\mathcal{F}$, удовлетворяет равенству $\langle \mathcal{F}^\dagger(g), f\rangle = \langle g, \mathcal{F}(f)\rangle$ для любых $g \in \mathbb{C}^G$ и $f \in \mathbb{C}^G$. Следовательно, для любых χ и ψ

$$\left\langle \left(\mathcal{F}^\dagger \mathcal{F}\right)(\chi), \psi \right\rangle = \left\langle \mathcal{F}(\chi), \mathcal{F}(\psi) \right\rangle = \left\langle \delta^\chi, \delta^\psi \right\rangle = \delta^\chi(\psi).$$

Отсюда следует, что $\mathcal{F}^{-1} = \mathcal{F}^\dagger$. Что до конкретной формулы, заметим, что

$$f(x) = \left(\mathcal{F}^{-1}\,\widehat{f}\right)(\delta_x) = \widehat{f}\left(\mathcal{F}\,\delta_x\right) = \sum_{\chi \in \widehat{G}} \chi(x)\,\widehat{f}(\chi).$$

Второе равенство следует из определения 10.15. □

10.2.3. Двойственность Понтрягина

Связь между G и $\hat{G}$ или между $\mathbb{C}^G$ и $\mathbb{C}^{\hat{G}}$ называется *двойственностью Понтрягина*. Эта двойственность переносится на подгруппы G и $\hat{G}$.

Определение 10.17 (аннулятор). Пусть G – конечная абелева группа. Аннулятором подмножества $H \subseteq G$ называется подгруппа

$$H^\perp = \{\, \chi \in \hat{G} \mid \forall x \in H : \chi(x) = 1 \,\}.$$

Аналогично для подгруппы $H \subseteq \hat{G}$ аннулятором H называется группа

$$H^\perp = \{\, x \in G \mid \forall \chi \in H : \chi(x) = 1\}.$$

Эти определения эквивалентны каноническому изоморфизму из теоремы 10.11.

Из определения 10.17 следует, что $H^{\perp\perp} = H$.

Если $\{0\} \subseteq H \subseteq K \subseteq G$, то $\{1\} \subseteq K^\perp \subseteq H^\perp \subseteq \hat{G}$. Иначе говоря, «взятие аннулятора» отображает подгруппы G в подгруппы $\hat{G}$ и наоборот, но изменяет направление включения на противоположное. Следующий результат дает более детальную характеристику групп $H^\perp$ и $K^\perp$.

Теорема 10.18. *Пусть H – подгруппа G. Существует изоморфизм между G/H и $\hat{H}^\perp$, определяемый передачей $x + H$ отображению вычисления* $\mathrm{ev}_x : H^\perp \to \mathbb{C}$. *Кроме того, этот изоморфизм приводит к следующему равенству подпространств:*

$$\mathrm{Span}\{\chi \mid \chi \in H^\perp\} = \mathrm{Span}\{f \circ \pi_H \mid f \in \mathbb{C}^{G/H}\},$$

где $\pi_H : G \to G/H$ – отображение проецирования $\pi_H(x) = x + H$.

Доказательство. Заметим, что отображение $\theta : G \to \hat{H}^\perp$, определенное как $\theta(x) = \mathrm{ev}_x$, является гомоморфизмом групп. По теореме 10.11, ev_x – характер $H^\perp$, определенный как $\chi \mapsto \chi(x)$. Следовательно, $\mathrm{ev}_x \equiv 1$ тогда и только тогда, когда x является элементом H. Отсюда следует, что ядро θ совпадает с H. Тогда первая теорема об изоморфизме групп показывает, что $x + H \mapsto \mathrm{ev}_x$ является изоморфизмом G/H и $\hat{H}^\perp$.

Равенство подпространств можно продемонстрировать следующим образом. Любая функция, принадлежащая оболочке $H^\perp$, постоянна на смежных классах H, потому что $\chi(x) = \chi(y)$ тогда и только тогда, когда $x/y \in H^\perp$. Следовательно:

$$\mathrm{Span}\{\chi \mid \chi \in H^1\} \subseteq \mathrm{Span}\{f \circ \pi_H \mid f \in \mathbb{C}^{G/H}\}.$$

Однако, в силу изоморфизма, размерности совпадают, и, значит, это включение является равенством. □

Теорема 10.18 упрощает вычисление преобразования Фурье функций, постоянных на смежных классах подгруппы $H \subseteq G$, т. е. имеющих вид $f \circ \pi_H$, где $f \in \mathbb{C}^{G/H}$. В анализе Фурье такие функции называются *периодическими*. По теореме 10.18, $f \circ \pi_H$ является линейной комбинацией характеров, принадлежащих H^1. Следовательно,

$$\widehat{f \circ \pi_H}(\chi) = \begin{cases} \widehat{f}(\chi_{G/H}), & \text{если } \chi \in H^1, \\ 0 & \text{в противном случае.} \end{cases}$$

В первом случае $\chi_{G/H}$ является характером G/H, полученным из χ посредством отображения $x + H \mapsto \chi(x)$. Он определен корректно, потому что $\chi \in H^1$. Действительно, по определению 10.14,

$$\begin{aligned} \widehat{f \circ \pi_H}(\chi) &= \frac{1}{|G|} \sum_{x \in G} \overline{\chi(x)} f(\pi_H(x)) \\ &= \frac{1}{|G/H|} \sum_{x+H \in G/H} \overline{\chi(x+H)} f(x+H) \\ &= \widehat{f}(\chi_{G/H}). \end{aligned}$$

Равенство подпространств в теореме 10.18 можно дуализировать, применив антиизоморфизм $x \mapsto x^*$ из теоремы 10.6 к обеим частям:

$$\mathrm{Span}\{\chi^* \mid \chi \in H^1\} = \mathrm{Span}\{\textstyle\sum_{g \in x+H} \delta_g \mid x + H \in G/H\}.$$

Существует вариант теоремы 10.18 для подгрупп $\hat{G}$; см. упражнение 10.10.

10.3. Историческая справка

Дополнительные сведения о линейной алгебре, обсуждаемой в этой главе, можно найти в большинстве учебников по линейной алгебре, например в книге Халмоша (1958). Теория анализа Фурье на конечных абелевых группах разрабатывается в нескольких книгах, в частности Терраса (1999).

Сведения о линейной алгебре из этой главы понадобятся в главе 11. Стимулом для изучения преобразования Фурье на произвольных абелевых группах, а не только на $\mathbb{F}_2^n$, служит тот факт, что в главе 11 заново выстраивается теория линейного криптоанализа в этой постановке – хотя это ни в коем случае не является основной целью главы. Исторически линейный криптоанализ впервые был обобщен на другие конечные абелевы группы в работе Беньера, Стерна и Воденэ. Более полную трактовку дал Бейн (2021).

10.4. Литература

Baignères, Thomas, Jacques Stern, and Serge Vaudenay (Aug. 2007). «Linear Cryptanalysis of Non Binary Ciphers». In: *SAC 2007*. Ed. by Carlisle M. Adams, Ali Miri,

and Michael J. Wiener. Vol. 4876. LNCS. Springer, Berlin, Heidelberg, pp. 184–211. doi: 10.1007/978-3-540-77360-3_13.

Beyne, Tim (Dec. 2021). «A Geometric Approach to Linear Cryptanalysis». In: *ASIACRYPT 2021, Part I*. Ed. by Mehdi Tibouchi and Huaxiong Wang. Vol. 13090. LNCS. Springer, Cham, pp. 36–66. doi: 10.1007/978-3-030- 92062-3_2.

Halmos, Paul R. (1958). *Finite-dimensional Vector Spaces*. 1st ed. Undergraduate Texts in Mathematics. Springer New York, NY.

Terras, Audrey (1999). *Fourier Analysis on Finite Groups and Applications*. London Mathematical Society Student Texts. Cambridge University Press, Cambridge.

10.5. Упражнения

Упражнение 10.1

Докажите, что функции δ_x, где $x \in G$, образуют базис $\mathbb{C}^G$:

1) покажите, что $\mathrm{Span}\{\delta_x \mid x \in G\} = \mathbb{C}^G$;
2) покажите, что функции δ^x, где $x \in G$, линейно независимы.

Упражнение 10.2

Пусть p и q – вещественные числа, большие 1, такие что $1/p + 1/q = 1$.

1. Покажите, что $xy \leqslant x^p/p + y^q/q$ для любых неотрицательных $x, y \in \mathbb{R}$.
2. Выведите отсюда, что $|f(g)| \leqslant \|f\|_q \|g\|_p$ для любых $f \in \mathbb{C}^G, g \in \mathbb{C}[G]$.
3. Покажите, что $\|f\|_p^\vee \leqslant \|f\|_q$.
4. Для любой $f \in \mathbb{C}^G$ постройте g такую, что $|f(g)| = \|f\|_q \|g\|_p$. Отсюда сделайте вывод, что $\|f\|_p^\vee = \|f\|_q$.

Упражнение 10.3

Докажите, что $\langle f, g\rangle = \langle f, g\rangle = \frac{1}{|G|}\sum_{x\in G} \overline{f(x)}\, g(x)$ определяет скалярное произведение на $\mathbb{C}^G$.

Упражнение 10.4

Цель этого упражнения – доказать теорему 10.6. Пусть $\theta : x \mapsto x^*$.

1. Докажите, что $\theta(x + y) = \theta(x) + \theta(y)$.
2. Докажите, что $\theta(\lambda x) = \lambda\theta(x)$.
3. Докажите, что θ обратимо. Что такое θ^{-1}?

Упражнение 10.5

Пусть V – подпространство конечномерного пространства со скалярным произведением U. Докажите, что:

1) $V^\perp$ является пространством со скалярным произведением;
2) ортогональное дополнение $V^\perp$ совпадает с самим V: $V^{\perp\perp} = V$;
3) $U = V \oplus V^\perp$.

Последний результат показывает, что ортогональное дополнение является алгебраическим дополнением.

Упражнение 10.6

Пусть V – подпространство конечномерного пространства со скалярным произведением U. Аннулятором V называется подпространство $V^0 \subseteq U^*$, определенное как

$$V^0 = \{u \in U^* \mid \forall v \in V : u(v) = 0\}.$$

Двойственно, аннулятором подпространства $W \subseteq U^*$ называется следующее подпространство U:

$$W^0 = \{u \in U \mid \forall w \in W : w(u) = 0\}.$$

Докажите, что

1) $\dim V + \dim V^0 = \dim U = \dim U^* = \dim W + \dim W^0$;
2) если U – пространство со скалярным произведением, то $(V^0)^* = V^\perp$ и $(W^0)^* = W^\perp$.

Во втором вопросе $x \mapsto x^*$ является антиизоморфизмом, в силу теоремы 10.6.

Упражнение 10.7

Понятие угла между векторами обобщается на подпространства конечномерного пространства с внутренним произведением W. Для подпространств U, $V \subseteq W$ определим линейное отображение $\langle V, U\rangle : U \to V$ как $\langle V, U\rangle = \pi_V \iota_U$, где ι_U: $U \to W$ – отображение включения, а π_V: $W \to V$ – ортогональная проекция на V.

1. Покажите, что если U и V – одномерные подпространства, натянутые на векторы u и v единичной нормы соответственно, то $\langle V, U\rangle : \lambda u \mapsto \langle v,u\rangle \lambda v$.
2. Покажите, что для любого вектора $u \in U$ никакой другой вектор V такой же длины не образует с u угол меньший, чем $\langle V, U\rangle(u)$.
3. Пусть $\sigma_1, \ldots, \sigma_d$ – сингулярные значения $\langle V, U\rangle$, соответствующие правым и левым сингулярным векторам $u_1, \ldots, u_d$ и $v_1, \ldots, v_d$ соответственно. Положим $U_i = U \cap \mathrm{Span}\{u_1, \ldots, u_{i-1}\}^\perp$ и $V_i = V \cap \mathrm{Span}\{v_1, \ldots, v_{i-1}\}^\perp$. Докажите, что для любого $i \in \{1, \ldots, d\}$

$$\sigma_i = \frac{\langle u_i, v_i\rangle}{\|u_i\|\,\|v_i\|} = \max_{\substack{u \in U_i \setminus \{0\} \\ v \in V_i \setminus \{0\}}} \frac{|\langle u, v\rangle|}{\|u\|\,\|v\|}.$$

Углы $0 \leqslant \theta_1 \leqslant \ldots \leqslant \theta_d \leqslant \pi/2$, такие что $\cos\theta_i = \sigma_i$, называются *главными углами* между подпространствами U и V. Сингулярные векторы – это направления, вдоль которых измеряются главные углы.

* Упражнение 10.8

Пусть $L : V \to V$ – линейное отображение на конечномерном пространстве со скалярным произведением V. Подпространство $U \subseteq V$ называется инвариантным подпространством L, если $L(U) \subseteq U$.

1. Докажите, что если U – инвариантное подпространство L, то $U^\perp$ – инвариантное подпространство $L^\dagger$ и, наоборот.
2. Воспользовавшись предыдущим результатом, докажите, что если отображение L самосопряженное, то существует ортонормированный базис V, такой что представление L в этом базисе является диагональной матрицей.

Упражнение 10.9

Докажите, что между $\hat{\hat{G}}$ и G существует канонический изоморфизм $x \mapsto \mathrm{ev}_x$, где $\mathrm{ev}_x : \chi \mapsto \chi(g)$ – отображение вычисления (см. теорему 10.11).

1. Покажите, что ev_x – характер $\hat{G}$ для любого $x \in G$.
2. Докажите, что $x \mapsto \mathrm{ev}_x$ – изоморфизм групп.

Упражнение 10.10

Цель этого упражнения – доказать аналог теоремы 10.18 для подгруппы $H \subseteq \hat{G}$. Для каждого вопроса используйте два разных рассуждения: одно, аналогичное доказательству теоремы 10.18, и другое, основанное на двойственности.

1. Укажите изоморфизм $\hat{G}/H$ в $\hat{H}^1$.
2. Покажите, что этот изоморфизм продолжается до следующего равенства подпространств:

$$\mathsf{Span}\left\{\delta_x \mid x \in H^1\right\} = \mathsf{Span}\left\{\sum_{\psi \in \chi H} \psi^* \mid \chi H \in \widehat{G}/H\right\}.$$

Глава 11

Геометрический подход

В этой главе мы заново – в последний раз – построим теорию линейного криптоанализа. Одна из причин, зачем это нужно, уже упоминалась в главе 9: существуют различные комбинаторные свойства, которые потенциально могут быть полезны, но для которых отсутствуют аналитические методы. Однако прежде чем приступать к этому вопросу, мы должны отступить назад и попытаться улучшить наше понимание линейного криптоанализа.

11.1. Геометрический взгляд

Пусть $\mathsf{F} : G \to H$ – криптографический примитив, например блочный шифр. Отправной точкой геометрического подхода служит формулировка криптоаналитических свойств F, таких как линейные аппроксимации, в терминах пары векторных пространств $\mathbb{C}[G]$ и $\mathbb{C}^H$. Этот взгляд весьма общий и с некоторыми модификациями применим также к другим важным методам, в т. ч. к дифференциальному и интегральному криптоанализам.

11.1.1. Криптоаналитические свойства

В простейшем случае векторные пространства одномерные, так что криптоаналитическое свойство определяется парой (u, v), где $u \in \mathbb{C}[G]$, а $v \in \mathbb{C}^H$. На интуитивном уровне смысл u и v следующий:

- вектор u представляет назначения весов (комплексных чисел) элементам G. Это способ отслеживать состояние набора входов или выходов;
- функция v отображает элементы G и – путем продолжения – $\mathbb{C}[G]$ в $\mathbb{C}$. Она представляет измерение или наблюдение состояния набора входов либо выходов.

Применение функции $\mathsf{F} : G \to H$ к состоянию преобразует назначение весов на G в соответствующее назначение на H. В простейшем случае, когда F – перестановка, это приводит к переупорядочению весов, которые были назначены элементам X. В разделе 11.1.2 описано действие функций F общего вида. Пока что достаточно будет сказать, что результат характеризуется линейной функцией $T^{\mathsf{F}} : \mathbb{C}[G] \to \mathbb{C}[H]$, которая отображает δ_x в $\delta_{\mathsf{F}(x)}$.

В криптоанализе редко бывает возможно вычислить точное состояние $T^{\mathsf{F}}u$. Однако достаточно вычислить $v(T^{\mathsf{F}}u)$. Следующее определение обобщает сказанное выше.

Определение 11.1 (криптоаналитическое свойство). Криптоаналитическим свойством функции $\mathsf{F} : G \to H$ называется пара (U, V), где U – подпространство $\mathbb{C}[G]$, а V – подпространство $\mathbb{C}^H$. Результат вычисления свойства в $u \in U$ и $v \in V$ равен $v(T^{\mathsf{F}}u)$.

Цель методов, с которыми мы познакомимся в этой главе, – оценить $v(T^{\mathsf{F}}u)$. Следующий пример криптоаналитического свойства полезно иметь в виду.

Пример 11.1. Пусть $\mathsf{F} : G \to H$ – функция, а X и Y – подмножества G и H соответственно. В этом примере определяется криптоаналитическое свойство, равное числу элементов $x \in X$, таких что $\mathsf{F}(x) \in Y$. Пусть δ_X – вектор, определенный как

$$\delta_X = \sum_{x \in X} \delta_x.$$

Конкретный пример для $G = \mathbb{F}_2^2$ показан на рис. 11.1. Определим $\delta^Y : H \to \mathbb{C}$ как

$$\delta^Y = \sum_{y \in Y} \delta^y.$$

То есть $\delta^Y(y) = 1$, если $y \in Y$, и $\delta^Y(y) = 0$ в противном случае. Линейное продолжение δ^Y на H дает линейный функционал, который суммирует свои входы по X. Следовательно, результатом вычисления свойства (U, V), где $U = \mathrm{Span}\{\delta_x\}$ и $V = \mathrm{Span}\{\delta^Y\}$, является

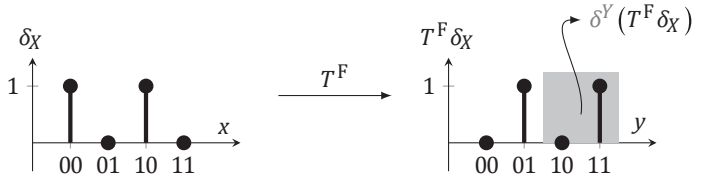


Рис. 11.1. Свойство $\mathsf{F}: \mathbb{F}_2^2 \to \mathbb{F}_2^2$ из примера 11.1, где $\mathsf{F}(x) = x + 11$, $X = \{00, 10\}$ и $Y = \{10, 11\}$. Результат вычисления $\delta^Y(T^{\mathsf{F}}\delta_X) = 1$

$$\delta^Y\left(T^{\mathsf{F}}\,\delta_X\right) = \sum_{x \in X} \delta^Y\left(\delta_{\mathsf{F}(x)}\right) = |\{x \in X \mid \mathsf{F}(x) \in Y\}|.$$

В первом равенстве используется тот факт, что T^{F} отображает δ_x в $\delta_{\mathsf{F}(x)}$. ▷

11.1.2. Распространение

Для функции $\mathsf{F} : G \to H$ общего вида линейное отображение T^{F} определяется следующим образом.

Определение 11.2 (прямой образ). Пусть $\mathsf{F} : G \to H$ – функция. Оператором прямого образа (pushforward) F называется линейное отображение $T^{\mathsf{F}} : \mathbb{C}[G] \to \mathbb{C}[H]$, определенное как

$$T^{\mathsf{F}}\delta_x = \delta_{\mathsf{F}(x)}$$

для любого $x \in G$.

Из определения 11.2 следует, что для вектора $u \in \mathbb{C}[G]$ с координатами u_x

$$T^{\mathsf{F}} u = \sum_{y \in H} \delta_y \sum_{\substack{x \in G \\ \mathsf{F}(x) = y}} u_x.$$

Матричное представление T^{F} относительно стандартных базисов $\mathbb{C}[G]$ и $\mathbb{C}[H]$ называется *матрицей переходов* F. Допуская некоторую вольность, она тоже обозначается T^{F}. Элементы матрицы переходов равны

$$T^{\mathsf{F}}_{y,x} = \begin{cases} 1, & \text{если} \quad y = \mathsf{F}(x), \\ 0 & \text{в противном случае.} \end{cases}$$

Существует двойственная версия определения 11.2, соответствующая матрице, получающейся транспонированием T^{F}. Оно описывает обратное распространение функции $v \in \mathbb{C}^H$ через F.

Определение 11.3 (обратный образ). Пусть $\mathsf{F} : G \to H$ – функция. Оператором обратного образа (pushback) F называется линейное отображение $T^{\mathsf{F}^\vee}: \mathbb{C}^H \to \mathbb{C}^G$, определенное как

$$T^{\mathsf{F}^\vee} \delta^y = \delta^y \circ \mathsf{F}$$

для любого $y \in H$.

Оператор обратного образа действительно является результатом транспонирования оператора прямого образа:

$$\left(T^{\mathsf{F}^\vee} \delta^y\right)(\delta_x) = \left(\delta^y \circ \mathsf{F}\right)(x) = \delta^y\left(\delta_{\mathsf{F}(x)}\right) = \delta^y\left(T^{\mathsf{F}} \delta_x\right).$$

Это также означает, что матричное представление $T^{\mathsf{F}^\vee}$ относительно стандартных базисов $\mathbb{C}^H$ и $\mathbb{C}^G$ описывается матрицей, являющейся результатом транспонирования F.

Пример 11.2. Матрица переходов S-блока для демонстрационного шифра из главы 1 равна следующей матрице перестановки:

$$T^{\mathsf{S}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Матричным представлением оператора обратного образа является результат транспонирования T^{S}. ▷

В следующей теореме формулируются два важных свойства операторов прямого образа. Все они имеют место также для операторов обратного образа, но в свойстве (2) порядок операций умножения следует изменить на противоположный.

Теорема 11.4. *Пусть* $\mathsf{F} : G \to H$ *– функция. Оператор прямого образа* T^{F} *функции* F *обладает следующими свойствами:*

(1) *Если* $\mathsf{F}((x_1, \ldots, x_n)) = (\mathsf{F}_1(x_1), \ldots, \mathsf{F}_n(x_n))$, *где* $\mathsf{F}_i : G_i \to H_i$ *такая, что* $G = \bigoplus_{i=1}^n G_i$ *и* $H = \bigoplus_{i=1}^n H_i$, *то* $T^{\mathsf{F}} = T^{\mathsf{F}_1} \otimes T^{\mathsf{F}_2} \otimes \ldots \otimes T^{\mathsf{F}_n}$.

(2) *Если* $\mathsf{F} = \mathsf{F}_r \circ \ldots \circ \mathsf{F}_2 \circ \mathsf{F}_1$, *то* $T^{\mathsf{F}} = T^{\mathsf{F}_r} \otimes \ldots \otimes T^{\mathsf{F}_2} \otimes T^{\mathsf{F}_1}$.

Доказательство. Достаточно доказать результат для $n = 2$ и $r = 2$. Из определения 11.2 и определения тензорных произведений в примере 10.5 следует, что

$$T^{\mathsf{F}} \underbrace{\delta_{(x_1, x_2)}}_{\delta_{x_1} \otimes \delta_{x_2}} = \underbrace{\delta_{(\mathsf{F}_1(x_1), \mathsf{F}_2(x_2))}}_{\delta_{\mathsf{F}_1(x_1)} \otimes \delta_{\mathsf{F}_2(x_2)}}.$$

Отсюда для любых $x_1 \in G_1$ и $x_2 \in G_2$ имеем

$$T^{\mathsf{F}} (\delta_{x_1} \otimes \delta_{x_2}) = T^{\mathsf{F}_1} \delta_{x_1} \otimes T^{\mathsf{F}_2} \delta_{x_2}.$$

Второе свойство также является прямым следствием определения 11.2:

$$T^{\mathsf{F}_2 \circ \mathsf{F}_1} \delta_x = \delta_{\mathsf{F}_2(\mathsf{F}_1(x))} = T^{\mathsf{F}_2} \delta_{\mathsf{F}_1(x)} = T^{\mathsf{F}_2} T^{\mathsf{F}_1} \delta_x,$$

и это справедливо для любого $x \in G$. □

11.1.3. Геометрия

В главе 10 векторные пространства $\mathbb{C}[G]$ и $\mathbb{C}^G$ были снабжены p-нормой. Как было отмечено, в частном случае $p = 2$ мы приходим к пространству со скалярным произведением. Оказывается, что евклидова норма играет важную роль в криптоанализе.

Если $(x_1, y_1), \ldots, (x_q, y_q)$ – выборка q пар (открытый текст, шифртекст), то несмещенная оценка $v(T^{\mathsf{F}} u)$ имеет вид

$$t = \frac{|G|}{q} \sum_{i=1}^{q} v(y_i)\, u_{x_i}.$$

Предположим, что случайные входы выбираются независимо и равномерно. Вариант простой модели постулирует, что для неправильных ключей выходы $\mathbf{y}_1, \ldots, \mathbf{y}_q$ являются независимыми и равномерно распределенными случайными величинами на H. Если, кроме того, $\sum_{x \in G} u_x = 0$ и $\sum_{y \in H} v(y) = 0$, то дисперсия[1] $\mathbf{t}$ для неправильных ключей равна

$$\frac{|G|^2}{q} \sum_{i=1}^{q} \mathbb{E}|u_{\mathbf{x}_i}|^2\, |v(\mathbf{y}_i)|^2 = \left(|G| \sum_{x \in G} |u_x|^2\right)\left(\frac{1}{|H|} \sum_{y \in H} |v(y)|^2\right) = \|u\|_2^2\, \|v\|_2^2,$$

где нормы определены, как в главе 10.

Словесно это означает, что произведение длин u и v является стандартным отклонением статистики критерия. При дополнительных предположениях, которые здесь подробно не обсуждаются, можно показать, что стандартное отклонение приблизительно одинаково для правильного ключа. Фактически

[1] Дисперсия комплексной случайной величины $\mathbf{z}$ равна $\mathbb{E}|\mathbf{z} - \mathbb{E}\mathbf{z}|^2$.

если мы допустим, что статистика критерия распределена нормально, то из результатов, приведенных в главах 4 и 7, следует, что информационная сложность обратно пропорциональна

$$\frac{|v(T^{\mathsf{F}}u)|}{\|u\|_2\,\|v\|_2}.$$

Однако к этому следует отнестись с долей скептицизма. Упомянутые выше предположения не всегда справедливы, поэтому 2-норма не всегда является правильной мерой длины. Грубо говоря, хотя 2-норма в некоторых случаях правильно улавливает локальную геометрию, глобальную геометрию она не определяет.

По большей части геометрический подход является комбинаторной теорией и не пытается решать статистические задачи, такие как определение информационной сложности данного свойства. Однако в некоторых случаях выбор конкретной нормы автоматически приводит к полезным статистическим результатам. Изучение объясняющих это причин завело бы нас за рамки современного состояния исследований.

Замечание 11.5. Как обсуждалось в главе 10, 2-норма индуцирована скалярным произведением. Это скалярное произведение ведет к изометрическому изоморфизму между $\mathbb{C}[G]$ и $\mathbb{C}^G$. В большей части настоящей главы этот изоморфизм не используется. Но при работе с 2-нормой мысленное представление криптоаналитических свойств как пары подпространств $\mathbb{C}[G]$ и $\mathbb{C}[H]$ (или $\mathbb{C}^G$ и $\mathbb{C}^H$) иногда помогает вырабатывать геометрическую интуицию.

11.2. Линейный криптоанализ

Операторы прямого и обратного образов теоретически позволяют вычислять криптоаналитические свойства (в смысле определения 11.1). Однако в стандартном базисе это непрактично. Большинство блочных шифров включают сложение с раундовыми ключами k, для которого оператор прямого образа по соглашению обозначается T^k. В идеале свойства, используемые при анализе, должны минимально зависеть от ключа.

Для $f \in \mathbb{C}^G$ функция $T^{k^\vee} f$ определяется как $x \mapsto f(x + k)$. Отсюда, как обсуждалось в главе 10, следует, что преобразование Фурье диагонализирует все матрицы $T^{k^\vee}$. Двойственно F диагонализирует матрицы T^k. Следовательно, имеет смысл брать преобразование Фурье всех криптоаналитических свойств.

11.2.1. Корреляционные матрицы

Рассмотрим одномерное криптоаналитическое свойство, определяемое векторами $u \in \mathbb{C}[G]$ и $v \in \mathbb{C}^H$. Преобразования Фурье u и v равны $\hat{u} = \mathrm{F}_G(u)$ и $\hat{v} = \mathcal{F}_H(v)$ соответственно. Вычисление свойства производится по формуле

$$v\left(T^{\mathsf{F}}\,u\right) = \left(\mathcal{F}_H^{-1}\,\widehat{v}\right)\left(T^{\mathsf{F}}\,\mathcal{F}_G^{-1}\,\widehat{u}\right) = \widehat{v}\left(\mathcal{F}_H\,T^{\mathsf{F}}\mathcal{F}_G^{-1}\,\widehat{u}\right).$$

Второе равенство следует из того, что $\mathrm{F}_H = \mathcal{F}_H^{-\vee}$ (см. стр. 142 в разделе 10.2.2). Отображение $\mathrm{F}_H T^{\mathsf{F}} \mathcal{F}_H^{-\vee}$ является преобразованием Фурье оператора прямого образа T^{F}.

Определение 11.6 (корреляционная матрица). Пусть $\mathsf{F} : G \to H$ – функция между коммутативными группами G и H. Определим $C^{\mathsf{F}} : \mathbb{C}[\hat{G}] \to \mathbb{C}[\hat{H}]$ как преобразование Фурье оператора прямого образа функции F:

$$C^{\mathsf{F}} = \mathcal{F}_H \, T^{\mathsf{F}} \, \mathcal{F}_G^{-1}.$$

Корреляционная матрица F является матричным представлением C^{F} относительно стандартных базисов $\mathbb{C}[\hat{G}]$ и $\mathbb{C}[\hat{H}]$.

Элементы корреляционной матрицы F равны

$$C^{\mathsf{F}}_{\chi,\psi} = \delta^{\chi}\left(C^{\mathsf{F}}\delta_{\psi}\right) = \chi\left(T^{\mathsf{F}}\psi^{*}\right) = \frac{1}{|G|}\sum_{x\in G}\chi\left(\mathsf{F}(x)\right)\overline{\psi(x)}.$$

Эти элементы являются результатами вычисления одномерных криптоаналитических свойств (u, v), где $u = \psi^*$ и $v = \chi$. Эквивалентно $\hat{u} = \delta_{\psi}$ и $\hat{v} = \delta^{\chi}$.

Пример 11.3. В линейном криптоанализе используются криптоаналитические свойства вида

$$(U, V) = (\mathrm{Span}\{\psi^*\}, \mathrm{Span}\{\chi\}).$$

Вычисление этого свойства для векторов с единичной нормой дает $C^{\mathsf{F}}_{\chi,\psi}$.

Если $G = \mathbb{F}_2^n$ и $H = \mathbb{F}_2^m$, то характеры группы ψ и χ равны

$$C^{\mathsf{F}}_{\chi,\psi} = \frac{1}{2^n}\sum_{x\in\mathbb{F}_2^n}(-1)^{u^{\mathsf{T}}x+v^{\mathsf{T}}\mathsf{F}(x)}.$$

С точностью до изоморфизма $u \mapsto (-1)^{u^{\mathsf{T}}x}$ между $\mathbb{F}_2^n$ и $\hat{\mathbb{F}}_2^n$ и $\mathbb{F}_2^m$ и $\hat{\mathbb{F}}_2^m$ элементы C^{F} можно индексировать буквами u и v вместо χ и ψ. Это возвращает нас к определению 2.3. ▷

Свойства корреляционных матриц, которые мы доказывали в главе 2 с помощью муторных вычислений, теперь оказываются прямыми следствиями теоремы 11.4. Объясняется это тем, что оператор прямого образа не зависит от выбора базиса. В частности, имеет место следующий результат.

Теорема 11.7. *Пусть $\mathsf{F} : G \to H$ – некоторая функция. Отображение C_{F} обладает следующими свойствами:*

(1) *Если* $\mathsf{F}((x_1, \ldots, x_n)) = (\mathsf{F}_1(x_1), \ldots, \mathsf{F}_n(x_n))$, где $\mathsf{F}_i : G_i \to H_i$ *такая, что* $G = \bigoplus_{i=1}^n G_i$ *и* $H = \bigoplus_{i=1}^n H_i$, *то* $C^{\mathsf{F}} = C^{\mathsf{F}_1} \otimes C^{\mathsf{F}_2} \otimes \ldots \otimes C^{\mathsf{F}_n}$.

(2) *Если* $\mathsf{F} = \mathsf{F}_r \circ \ldots \circ \mathsf{F}_2 \circ \mathsf{F}_1$, *то* $C^{\mathsf{F}} = C^{\mathsf{F}_r} \otimes \ldots \otimes C^{\mathsf{F}_2} \otimes C^{\mathsf{F}_1}$.

Доказательство. Как и в случае теоремы 11.4, достаточно доказать результат для $n = 2$ и $r = 2$. Первое свойство опирается на тот факт, что $\mathrm{F}_{G_1\oplus G_2} = \mathrm{F}_{G_1} \otimes \mathrm{F}_{G_2}$. Это следствие из теоремы 10.12. А именно

$$C^{\mathsf{F}} = \left(\mathcal{F}_{H_1}\otimes\mathcal{F}_{H_2}\right)\left(T^{\mathsf{F}_1}\otimes T^{\mathsf{F}_2}\right)\left(\mathcal{F}_{G_1}\otimes\mathcal{F}_{G_2}\right)^{-1} = \underbrace{\mathcal{F}_{H_1}\,T^{\mathsf{F}_1}\,\mathcal{F}_{G_1}^{-1}}_{C^{\mathsf{F}_1}} \otimes \underbrace{\mathcal{F}_{H_2}\,T^{\mathsf{F}_2}\,\mathcal{F}_{G_2}^{-1}}_{C^{\mathsf{F}_2}}.$$

Второе свойство явно вытекает из теоремы 11.4 (2) следующим образом:

$$C^{\mathsf{F}_2} C^{\mathsf{F}_1} = \left(\mathcal{F}_H\, T^{\mathsf{F}_2}\, \mathcal{F}_K^{-1}\right)\left(\mathcal{F}_K\, T^{\mathsf{F}_1}\, \mathcal{F}_G^{-1}\right) = \mathcal{F}_H\, T^{\mathsf{F}}\, \mathcal{F}_G^{-1} = C^{\mathsf{F}},$$

где F_K – преобразование Фурье для промежуточной группы K. □

Теорема 2.6 обобщается следующим образом.

Теорема 11.8. *Пусть* $\mathsf{F} : G \to H$ *определена как* $\mathsf{F}(x) = \mathsf{L}(x) + t$, *где* $\mathsf{L} : G \to H$ *– гомоморфизм групп, а* $t \in H$. *Корреляционная матрица* F *удовлетворяет равенству*

$$C^{\mathsf{F}}_{\chi,\psi} = \chi(t)\, \delta^{\chi \circ \mathsf{L}}(\psi).$$

Доказательство. Матрица C^t диагональная, на ее диагонали находятся собственные значения $\chi(t)$. По определению C^{L}, имеет место равенство

$$C^{\mathsf{L}}_{\chi,\psi} = (\chi \circ \mathsf{L})(\psi^*) = \langle \psi, \chi \circ \mathsf{L} \rangle = \delta^{\chi \circ \mathsf{L}}(\psi).$$

Последнее равенство справедливо в силу ортогональности характеров (теорема 10.13). Окончательный результат следует из того, что $C^{\mathsf{F}} = C^t C^{\mathsf{L}}$. □

11.2.2. Множественный линейный криптоанализ

Множественный линейный криптоанализ опирается на криптоаналитические свойства (U, V), где

$$U = \mathrm{Span}\{\psi_1^*, \psi_2^*, \ldots, \psi_n^*\},$$
$$V = \mathrm{Span}\{\chi_1, \chi_2, \ldots, \chi_m\}.$$

В главе 6 было показано, что многомерные линейные аппроксимации специальны, потому что их корреляции характеризуют распределение вероятностей линейной проекции пар (открытый текст, шифртекст) для случайных равномерно распределенных входов.

Многомерный линейный криптоанализ обобщается на произвольные конечные абелевы группы как частный случай множественного линейного криптоанализа, где $X_1 = \{\psi_1, \ldots, \psi_n\}$ и $Y_1 = \{\chi_1, \ldots, \chi_m\}$ – подгруппы G и H соответственно.

Обозначим $\pi_Y : H \to H/Y$ проекцию, определенную как $\pi_Y(h) = h + Y$. По теореме 10.18, подпространство $V = \mathrm{Span}\,\{\chi \mid \chi \in Y^{\perp}\}$ обладает следующим свойством:

$$V = \mathrm{Span}\left\{f \circ \pi_Y \mid f \in \mathbb{C}^{H/Y}\right\} = \mathrm{Span}\left\{\delta^{h+Y} \mid h + Y \in H/Y\right\},$$

где $\delta^{h+Y} = \sum_{y \in h+Y} \delta^y$, как в примере 11.1. Как было сказано в конце раздела 10.2.3, аналогичное равенство имеет место для подпространства $U = \mathrm{Span}\,\{\psi^* \mid \psi \in X^{\perp}\}$. Оно получается применением антиизоморфизма $x \mapsto x^*$ из теоремы 10.6:

$$U = \mathrm{Span}\,\{\delta_{g+X} \mid g + X \in G/X\}.$$

Результатом вычисления (U, V) для $\delta_{g+X}/|X| \in U$ и $\delta^{h+Y} \in V$ является

$$\frac{1}{|X|} \sum_{x \in g+X} \delta^{h+Y}(\mathsf{F}(x) + Y) = \Pr_{\mathbf{x}}\left[\mathsf{F}(\mathbf{x}) \equiv h \bmod Y\right],$$

где $\mathbf{x}$ – случайная равномерно распределенная величина на смежном классе $g+X$. Это та же самая вероятность, что в следствии 6.6. Из равенств указанных выше подпространств следует, что эту вероятность можно выразить в терминах корреляций линейных аппроксимаций с характерами, принадлежащими $X^\perp$ и $Y^\perp$. В упражнении 11.4 вам будет предложено сделать это явно.

11.3. Точное распространение

В первой части главы 9 обсуждались «точные» обобщения линейного криптоанализа. Примеры включают свойства насыщения, линейные аппроксимации с нулевой корреляцией и инварианты.

11.3.1. Прямое распространение

Под распространением подпространства $U \subseteq \mathbb{C}[G]$ посредством функции $\mathsf{F} : G \to H$ подразумевается, что нужно определить образ $T^{\mathsf{F}}U$. В общем случае это практически неосуществимо, потому-то вместо этого и используются криптоаналитические свойства, но в некоторых случаях можно показать, что $T^{\mathsf{F}}U \subseteq W$, где W – подпространство $\mathbb{C}[H]$.

Пример 11.4. Если $\mathsf{F}(X) \subseteq Y$, где $X \subseteq G$ и $Y \subseteq H$, то

$$T^{\mathsf{F}}\,\mathrm{Span}\,\{\delta_x \mid x \in X\} \subseteq \mathrm{Span}\,\{\delta_y \mid y \in Y\}.$$

Из этого включения следует, что результатом вычисления криптоаналитического свойства (U, V), где $U = \mathrm{Span}\{\delta_X\}$ и $V = \mathrm{Span}\{\delta^Y\}$, является $\delta^Y(T^{\mathsf{F}}\delta_X) = |X|$. ▷

Не каждому точному распространению соответствует включение множеств.

Пример 11.5. Описанный в разделе 9.1.1 метод нахождения свойств насыщения основан на описании состояния как одного из нескольких возможных множеств без указания на то, что происходит с отдельными элементами этих множеств. То есть для семейств множеств $S_1, \ldots, S_n$ и $T_1, \ldots, T_m$, где $m \geqslant n$,

$$T^{\mathsf{F}}\,\mathrm{Span}\,\{\delta_{S_1}, \ldots, \delta_{S_n}\} \subseteq \mathrm{Span}\,\{\delta_{T_1}, \ldots, \delta_{T_m}\}.$$

Например, для «функции проецирования» $\mathsf{P} : H \to Y$ (см. раздел 9.2.3) множества $T_1, \ldots, T_m$ можно охарактеризовать свойством $T^{\mathsf{P}}\delta_{T_i} \in \mathrm{Span}\{\delta_Y\}$. Пусть (U,V) – свойство, для которого $U = \mathrm{Span}\{\delta_{S_i}\}$ и $V = \mathrm{Span}\{\delta^y \circ \mathsf{P} \mid y \in Y\}$. В силу включения выше, вычисление свойства (U, V) дает

$$T^{\mathsf{F}}\,\mathrm{Span}\{\delta_x \mid x \in X\} \subseteq \mathrm{Span}\{\delta_y \mid y \in Y\}.$$

То есть $\mathsf{P} \circ \mathsf{F}$ насыщено на входном множестве S_i. ▷

11.3.2. Обратное распространение

В дополнение к распространению подпространства $\mathbb{C}[G]$ «вперед», как описано в разделе 11.3.1, можно распространить подпространство $V \subseteq \mathbb{C}^H$ назад посредством функции $\mathsf{F} : G \to H$. Для этого нужно показать, что $T^{\mathsf{F}^\vee} V \subseteq W$, где W – подпространство $\mathbb{C}^G$.

Пример 11.6. Пусть $f : G \to X$ и $g : H \to Y$ – функции. Этот результат является обобщением понятия нелинейной аппроксимации из раздела 9.2.2. Если

$$T^{\mathsf{F}^\vee} \operatorname{Span}\{\delta^y \circ g \mid y \in Y\} \subseteq \operatorname{Span}\{\delta^x \circ f \mid x \in X\},$$

то существует функция $h : X \to Y$, такая что $h \circ f = g \circ \mathsf{F}$. ▷

11.3.3. Нулевая корреляция

Свойство (U, V), вычисление которого дает нуль для всех $u \in U$ и $v \in V$, называется свойством с нулевой корреляцией. Пусть $\mathsf{F} = \mathsf{F}_2 \circ \mathsf{F}_1$, где $\mathsf{F}_1 : G \to H$ и $\mathsf{F}_2 : H \to K$. Аннулятор подпространства $W \subseteq \mathbb{C}^H$ определяется как (см. упражнение 10.6)

$$W^0 = \{x \in \mathbb{C}[H] \mid \forall w \in W : w(x) = 0\}.$$

Если $T^{\mathsf{F}_1} U \subseteq W^0$ и $T^{\mathsf{F}_2^\vee} V \subseteq W$ для некоторого подпространства W, то (U, V) является свойством с нулевой корреляцией:

$$v\left(T^{\mathsf{F}} u\right) = \left(T^{\mathsf{F}_2^\vee} v\right)\left(T^{\mathsf{F}_1} u\right) = 0$$

для любых $u \in U$ и $v \in V$. Этот обобщает принцип потери посередине из раздела 8.2, который использовался для нахождения линейных аппроксимаций с нулевой корреляцией.

Пример 11.7. Пусть $S_1, \ldots, S_n$ и $T_1, \ldots, T_m$ – два семейства множеств, как в примере 11.5, так что $T^{\mathsf{P}} \delta_{T_i} \in \operatorname{Span}\{\delta_Y\}$ для функции $\mathsf{P} : H \to Y$. Как и раньше, предположим, что $T^{\mathsf{F}} \operatorname{Span}\{\delta_{S_1}, \ldots, \delta_{S_n}\} \subseteq \operatorname{Span}\{\delta_{T_1}, \ldots, \delta_{T_m}\}$. Пусть $U = \operatorname{Span}\{\delta_{S_i}\}$ и для некоторой произвольной константы $z \in Y$

$$V = \operatorname{Span}\{\delta^y \circ \mathsf{P} - \delta^z \circ \mathsf{P} \mid y \in Y\}.$$

Так как $V^0 \supseteq \operatorname{Span}\{\delta_{T_1}, \ldots, \delta_{T_m}\}$, свойство (U,V) является свойством с нулевой корреляцией. ▷

11.3.4. Инварианты

Инвариантом функции $\mathsf{F} : G \to G$ называется подпространство $U \subseteq \mathbb{C}[G]$, такое что $T^{\mathsf{F}} U \subseteq U$.

Пример 11.8. Напомним (см. главу 9), что нелинейным инвариантом функции $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^n$ называется функция $f : \mathbb{F}_2^n \to \mathbb{F}_2$ такая, что существует постоянная c, такая что $f(\mathsf{F}(x)) = f(x) + c$ для любого $x \in \mathbb{F}_2^n$. Эквивалентно, если $S = \{x \in \mathbb{F}_2^n \mid f(x) = 1\}$, то либо $\mathsf{F}(S) \subseteq S$ и $\mathsf{F}(\mathbb{F}_2^n \setminus S) \subseteq \mathbb{F}_2^n \setminus S$, либо $\mathsf{F}(S) \subseteq \mathbb{F}_2^n \setminus S$ и $\mathsf{F}(\mathbb{F}_2^n \setminus S) \subseteq S$. Этому инварианту соответствует следующее векторное пространство U:

$$U = \operatorname{Span}\left\{\delta_S, \delta_{\mathbb{F}_2^n \setminus S}\right\}.$$

Альтернативно следующее подпространство $V \subseteq \mathrm{C}_G$ обладает свойством $T^{\mathsf{F}^\vee} V \subseteq V$:

$$V = \operatorname{Span}\{\delta^0 \circ f, \delta^1 \circ f\}.$$

Пространство V состоит из всех комплексных функций на $\mathbb{F}_2$, для которых обратный образ посредством f совпадает с $\mathbb{F}_2^n$. Чтобы проверить выполнение инварианта, можно вычислить свойство (U, V). ▷

Чтобы различить пространства U и V в примере 11.8, U можно назвать прямым инвариантом, а V – обратным инвариантом. Проверьте, что если U – прямой инвариант, то U^0 – обратный инвариант и наоборот. Кроме того, с помощью антиизоморфизма $x \mapsto x^*$ из теоремы 10.6 можно показать, что U является прямым инвариантом тогда и только тогда, когда U^* – обратный инвариант.

Инварианты связаны с собственными векторами T^{F}.

Теорема 11.9. *Любой инвариант U перестановки $\mathsf{F} : G \to G$ имеет базис, состоящий из собственных векторов T^{F}.*

Доказательство. Оператор прямого образа T^{F} допускает диагонализацию. Действительно, поскольку F^n является тождественной функцией для некоторого $n \geqslant 1$, минимальный многочлен T^{F} делит $x^n - 1$. Этот многочлен имеет различные корни в поле $\mathbb{C}$.

Если U является инвариантом, а F – перестановкой, то $T^{\mathsf{F}}U = U$. Отсюда следует, что минимальный многочлен ограничения $T^{\mathsf{F}}|_U : U \to U$ делит минимальный многочлен T^{F}. Следовательно, $T^{\mathsf{F}}|_U$ допускает диагонализацию. □

Из теоремы 11.9 следует, что преобразование Фурье инварианта имеет базис, состоящий из собственных векторов C^{F}.

Пример 11.9. Инвариант U из примера 11.8 натянут на два собственных вектора T^{F}:

$$U = \mathsf{Span}\left\{\delta_S + \delta_{\mathbb{F}_2^n \setminus S},\ \delta_S - \delta_{\mathbb{F}_2^n \setminus S}\right\}.$$

Первый вектор равен $\delta_S + \delta_{\mathbb{F}_2^n \setminus S} = \delta_{\mathbb{F}_2^n}$. Фактически это собственный вектор T^{F} для любой перестановки F. Аналогично V – оболочка, натянутая на два собственных вектора $T^{\mathsf{F}^\vee}$:

$$V = \mathrm{Span}\,\{x \mapsto 1,\ x \mapsto (-1)^{f(x)}\}.$$

В частности, преобразование Фурье $(-1)^f$ является собственным вектором $C^{\mathsf{F}^\vee}$. ▷

11.4. Приближенное распространение

Вычисление любого криптоаналитического свойства определяет линейную функцию, называемую его отображением аппроксимации. Это отображение используется, чтобы склеить последовательность свойств функций в свойство их композиции.

11.4.1. Отображения аппроксимации

Данные всех вычислений $\mathrm{v}(T^{\mathsf{F}}u)$ криптоаналитического свойства (U, V) эквивалентны линейному отображению $U \to \mathbb{C}[H]/V^0$, определенному как

$$u \mapsto T^{\mathsf{F}}u + V^0.$$

Здесь V^0 – аннулятор подпространства V. То есть состояние известно только с точностью до сложения с вектором, который не может быть «измерен» с помощью функций, принадлежащих V.

Описанное выше отображение можно было бы также назвать отображением аппроксимации (U, V), но проблема в том, что для того чтобы иметь возможность построить композицию отображения $U \to \mathbb{C}[H]/V^0$ с отображением другого свойства, необходимо вложение $\mathbb{C}[H]/V^0$ в $\mathbb{C}[H]$. К счастью, выбор алгебраического дополнения V решает эту проблему. Далее в данной главе используются ортогональные дополнения:[1] $\mathbb{C}^H = V \oplus V^\perp$. В силу упражнения 10.6 и обозначив $x \mapsto x^*$ антиизоморфизм из теоремы 10.6, имеем $(V^\perp)^0 = V^*$. Отсюда, поскольку $V^0 \cap (V^\perp)^0 = (V \oplus V^\perp)^0 = \{0\}$, имеем

$$\mathbb{C}[H] = V^0 \oplus V^*.$$

Смысл всего этого в том, что $\mathbb{C}[H]/V^0$ можно идентифицировать с помощью V^*, чтобы получить отображение $U \to V^*$. Важно, что V^* является подпространством $\mathbb{C}[H]$. Имейте в виду, что в следующем далее определении и обсуждении свойством является (U, V^*), а не (U, V), чтобы упростить обозначения. Это разумно, т. к. с любой точки зрения $V^{**} = V$.

Определение 11.10. Отображением аппроксимации криптоаналитического свойства (U, V^*) функции $\mathsf{F}: G \to H$ называется линейное отображение $\langle V, U\rangle_\mathsf{F} : U \to V$, определенное как

$$\langle V, U\rangle_\mathsf{F} = \pi_V T^\mathsf{F} \iota_U,$$

где $\iota_U : U \to \mathbb{C}[G]$ – включение, а $\pi_V : \mathbb{C}[H] \to V$ – ортогональная проекция.

Идея отображения аппроксимации состоит в том, что оно преобразует $u \in U$ в вектор, принадлежащий V, который аппроксимирует $T^\mathsf{F}u$ в следующем смысле. Для любых $u \in U$ и $v \in V^*$

$$v\langle V, U\rangle_\mathsf{F} u = v(\pi_V T^\mathsf{F} \iota_U u) = v(T^\mathsf{F}u).$$

Последнее равенство следует из того, что $\iota_U u = u$ и $T^\mathsf{F}u = \pi_V T^\mathsf{F}u + \pi_{V^\perp} T^\mathsf{F}u$, в сочетании с тем фактом, что V^* является аннулятором $V^\perp$. Иначе говоря, замена T^F его отображением аппроксимации не влияет на вычисление свойства.

11.4.2. Геометрия

По теореме о наилучшей аппроксимации (теорема 10.8), аппроксимации, полученные применением $\langle V, U\rangle_\mathsf{F}$, являются геометрически наилучшими из возможных. Качество криптоаналитического свойства измеряется его *главными корреляциями*.

Определение 11.11 (главные корреляции). Пусть (U, V^*) – криптоаналитическое свойство функции $\mathsf{F}: G \to H$. Положим $d = \min\{\dim U, \dim V\}$. Главными корреляциями свойства (U, V^*) называются d наибольших сингулярных значений его отображения аппроксимации $\langle V, U\rangle_\mathsf{F}$.

[1] Этого «произвольного» выбора можно избежать, воспользовавшись небольшим обобщением определения 11.1.

Если F инъективна, то главные корреляции (U, V^*) равны косинусам d наименьших главных углов между подпространствами $T^{\mathsf{F}}U$ и V (см. упражнение 10.7). Главная корреляция линейной аппроксимации $(\mathrm{Span}\{\psi^*\}, \mathrm{Span}\{\chi\})$ равна абсолютной величине ее корреляции.

У главных корреляций имеется также статистическая интерпретация. Не вдаваясь в детали, скажем, что если $\sigma_1, \ldots, \sigma_r$ – первые r главных корреляций некоторого свойства, то (в предположениях, которые мы здесь опускаем) минимальная информационная сложность проверки гипотезы на основе оценок с известным открытым текстом не более r результатов вычисления этого свойства обратно пропорциональна

$$\sum_{i=1}^{r} \sigma_i^2.$$

Сумма квадратов всех главных корреляций множественной линейной аппроксимации равна ее емкости.

11.4.3. Принцип доминирующих следов

Предположим, что $\mathsf{F} = \mathsf{F}_r \circ \ldots \circ \mathsf{F}_1$, где $\mathsf{F}_i : G_i \to G_{i+1}$. Для линейного криптоанализа свойство умножения корреляционных матриц приводит к понятию линейных следов $(\chi_1, \ldots, \chi_{r+1})$:

$$C^{\mathsf{F}}_{\chi_{r+1}, \chi_1} = \sum_{\chi_2, \ldots, \chi_r} \prod_{i=1}^{r} C^{\mathsf{F}_i}_{\chi_{i+1}, \chi_i}.$$

В этой формуле корреляции $C^{\mathsf{F}}_{\chi_{i+1}, \chi_i}$ являются результатами вычисления криптоаналитических свойств функций $\mathsf{F}_1, \ldots, \mathsf{F}_r$.

Следующий результат показывает, что аналогичное выражение существует для отображения аппроксимации произвольных свойств F. Последовательность векторных пространств $(U_1, U_2, \ldots, U_{r+1})$, или, эквивалентно, совместимых криптоаналитических свойств $(U_1, U_2^*), (U_2, U_3^*), \ldots, (U_r, U_{r+1}^*)$, называется следом.

Теорема 11.12. *Для* $1 \leqslant i \leqslant r + 1$ *обозначим* Ω_i *множество ортогональных подпространств* $\mathbb{C}[G_i]$, *такое что* $\mathbb{C}[G_i] = \bigoplus_{U \in \Omega^i} U$. *Для любого свойства* (U_1, U_{r+1}^*) *функции* F, где $U_1 \in \Omega_1$ *и* $U_{r+1} \in \Omega_{r+1}$,

$$\langle U_{r+1}, U_1 \rangle_{\mathsf{F}} = \sum_{U_2, \ldots, U_r} \langle U_{r+1}, U_r \rangle_{\mathsf{F}_r} \cdots \langle U_3, U_2 \rangle_{\mathsf{F}_2} \langle U_2, U_1 \rangle_{\mathsf{F}_1},$$

где суммирование производится по всем $(U_2, \ldots, U_r) \in \prod_{i=2}^{r} \Omega_i$.

Доказательство. По определению, $\langle U_{r+1}, U_i \rangle_{\mathsf{F}_r \circ \cdots \circ \mathsf{F}_i} = \pi_{U^{r+1}} T^{\mathsf{F}_r \circ \cdots \circ \mathsf{F}_i} \iota_{U_i}$. Кроме того, по определению Ω_{i+1}, отображение $\sum_{U \in \Omega_{i+1}} \pi_U$ тождественно. Отсюда

$$\langle U_{r+1}, U_i \rangle_{\mathsf{F}_r \circ \cdots \circ \mathsf{F}_i} = \sum_{U_{i+1} \in \Omega_{i+1}} \langle U_{r+1}, U_{i+1} \rangle_{\mathsf{F}_r \circ \cdots \circ \mathsf{F}_{i+1}} \langle U_{i+1}, U_i \rangle_{\mathsf{F}_i}.$$

Результат получается повторным применением этого равенства для $i = 1, \ldots, r - 1$. □

Поскольку знание отображения аппроксимации некоторого свойства эквивалентно знанию всех результатов его вычисления, теорема 11.12 дает способ склеить свойства $\mathsf{F}_1, \ldots, \mathsf{F}_r$. На практике суммирование по всем следам аппроксимируется суммированием по небольшому множеству доминирующих следов.

11.5. Историческая справка

Отправной точкой геометрического подхода стала идея о том, что корреляционные матрицы представляют линейные отображения. Если отнестись к этой точке зрения всерьез, то важно понимать, на каких векторных пространствах эти отображения действуют. Первым применением этого подхода стал анализ инвариантов в работе Бейна (2018). Линейный криптоанализ и его обобщения обсуждались в работе Бейна (2021).

Случай линейного криптоанализа служит введением в геометрический подход вообще. Он важен для понимания других криптоаналитических методов, таких как дифференциальный и интегральный криптоанализы, и связей между ними (Бейн, 2023).

11.6. Литература

Beyne, Tim (Dec. 2018). «Block Cipher Invariants as Eigenvectors of Correlation Matrices». In: *ASIACRYPT 2018, Part I*. Ed. by Thomas Peyrin and Steven Galbraith. Vol. 11272. LNCS. Springer, Cham, pp. 3–31. doi: 10.1007/978-3-030-03326-2_1.

Beyne, Tim (Dec. 2021). «A Geometric Approach to Linear Cryptanalysis». In: *ASIACRYPT 2021, Part I*. Ed. by Mehdi Tibouchi and HuaxiongWang. Vol. 13090. LNCS. Springer, Cham, pp. 36–66. doi: 10.1007/978-3-030-92062-3_2.

Beyne, Tim (June 2023). «A Geometric Approach to Symmetric-key Cryptanalysis». PhD thesis. KU Leuven.

11.7. Упражнения

Упражнение 11.1

Пусть G – конечная абелева группа. Неподвижной точкой функции $\mathsf{F} : G \to G$ называется элемент $x \in G$, такой что $\mathsf{F}(x) = x$. Напомним, что следом $\mathrm{Tr}\, A$ матрицы A называется сумма элементов на ее главной диагонали. Докажите, что $\mathrm{Tr}\, C^{\mathsf{F}}$ равен количеству неподвижных точек F.

Упражнение 11.2

Пусть $\mathsf{F} : G \to H$ – некоторая функция. Пара входов (x, y) называется коллизией для F, если $\mathsf{F}(x) = \mathsf{F}(y)$. Докажите следующую формулу для вероятности того, что случайная равномерно распределенная пара входов является коллизией:

$$|H| \Pr_{\mathbf{x},\mathbf{y}}\left[\mathsf{F}(\mathbf{x}) = \mathsf{F}(\mathbf{y})\right] = \sum_{\chi \in \widehat{H}} \left|C^{\mathsf{F}}_{\chi,1}\right|^2.$$

В формуле выше $\mathbf{x}$ и $\mathbf{y}$ – случайные равномерно распределенные величины на G, а C^{F} – корреляционная матрица F.

Упражнение 11.3

Пусть $\mathbb{F}_q$ – конечное поле порядка q, а f – полином над $\mathbb{F}_q$ степени $d \geqslant 2$, взаимно простой с q. Одним из следствий гипотезы Римана для кривых над конечными полями является следующая оценка экспоненциальной суммы:

$$\left| \sum_{x \in \mathbb{F}_q} \exp\left(\frac{2\pi i \operatorname{Tr} f(x)}{p} \right) \right| \leq (d-1)\sqrt{q}.$$

Этот результат называется границей Вейля.

1. Пусть $\mathsf{F} : \mathbb{F}_q \mapsto \mathbb{F}_q$ – кубическая функция, определенная как $\mathsf{F}(x) = x^3$. Докажите, что

   $$\left| C^{\mathsf{F}}_{\chi,\psi} \right| \leq 2/\sqrt{q}$$

   для любого нетривиального характера χ в предположении, что q не является степенью тройки.

2. Предположим, что $q \equiv 2 \pmod 3$. Пусть $\mathsf{G} : \mathbb{F}_q \mapsto \mathbb{F}_q$ – функция, определенная как $\mathsf{G}(x) = (x^3 + k)^{1/3}$. Докажите, что если $k \neq 0$ и q нечетно, то

   $$\left| C^{\mathsf{G}}_{\chi,\psi} \right| \leq 2/\sqrt{q}$$

   для любого нетривиального характера χ.

3. Почему во втором вопросе возникает проблема, если q четно?

Упражнение 11.4

Пусть G и H – конечные абелевы группы с подгруппами X и Y соответственно, как в разделе 11.2.2. Пусть $\mathsf{F} : G \to H$ – некоторая функция, и рассмотрим вероятности

$$\Pr_{\mathbf{x}}\left[\mathsf{F}(\mathbf{x}) \equiv h \bmod Y\right],$$

где $\mathbf{x}$ – случайная равномерно распределенная величина на смежном классе $g + X$. По теореме 10.18, эти вероятности можно выразить в виде линейных комбинаций корреляций $C^{\mathsf{F}}_{\chi,\psi}$, где $\psi \in X^1$, $\chi \in Y^1$.

1. Докажите следующее неравенство для случайной величины $\mathbf{x}$, равномерно распределенной на $g + X$:

   $$\Pr_{\mathbf{x}}\left[\mathsf{F}(\mathbf{x}) \equiv h \bmod Y\right] = \frac{|Y|}{|H|} \sum_{\substack{\psi \in X^1 \\ \chi \in Y^1}} \overline{\chi(h)}\, \psi(g)\, C^{\mathsf{F}}_{\chi,\psi}.$$

   Избегайте громоздких вычислений типа тех, что встретились в доказательстве теоремы 6.3.

2. Докажите обратное соотношение: запишите $C^{\mathsf{F}}_{\chi,\psi}$ (где $\psi \in X^1, \chi \in Y^1$) в виде линейной комбинации вероятностей для всех смежных классов $g + X$ и $h + Y$.

Упражнение 11.5

Приведите пример функции $\mathsf{F} : G \to G$, такой что T^{F} не допускает диагонализации.

* Упражнение 11.6

Пусть G – конечная абелева группа, а $\mathsf{F} : G \to G$ – перестановка. Предположим, что F имеет ℓ непересекающихся циклов длин $l_1, \ldots, l_\ell$, где i-й цикл содержит значения $(x_{i,1}, x_{i,2}, \ldots, x_{i,l_i})$.

1. Чему равны собственные значения T^{F}?
2. Выпишите соответствующие собственные векторы T^{F} и C^{F}.
3. Опишите перестановку $\mathsf{F} : \mathbb{F}_2^{1337} \to \mathbb{F}_2^{1337}$, не имеющую нетривиальных инвариантных подпространств.

Упражнение 11.7

Пусть $\mathsf{P}_1 : G \to X$ и $\mathsf{P}_2 : H \to Y$ – сбалансированные «проекции», а $\mathsf{F} : G \to H$ – некоторая функция. Заметим, что P_1 индуцирует следующее разбиение G:

$$G = \bigcup_{x \in X} \mathsf{P}_1^{-1}(x).$$

Далее мы полагаем $u = \delta_{\mathsf{P}_1^{-1}(x)}$ и $v = \delta^y \circ \mathsf{P}_2$ для $x \in X$ и $y \in Y$.

1. Докажите следующее равенство:

$$v(T^{\mathsf{F}} u) = |\{z \in G \mid \mathsf{P}_1(z) = x \wedge \mathsf{P}_2(\mathsf{F}(z)) = y\}|$$

для любых $x \in X$, $y \in Y$.

2. Выведите отсюда, что координаты отображения аппроксимации $(\mathrm{Span}\{u\}, \mathrm{Span}\{v\})$ в нормированных базисах имеют вид

$$\Pr_{\mathbf{z}}[\mathsf{P}_2(\mathsf{F}(\mathbf{z})) = y \mid \mathsf{P}_1(\mathbf{z}) = x],$$

где вероятность вычисляется для случайной равномерно распределенной величины $\mathbf{z}$.

3. Предположим, что $G = \mathbb{F}_2^n$, $X = Y = \mathbb{F}_2$, $\mathsf{P}_1(x) = u^{\mathsf{T}}x$ и $\mathsf{P}_2(x) = v^{\mathsf{T}}x$. Покажите, что если F является перестановкой, то существуют такие базисы, что матричное представление $\langle V, U\rangle_{\mathsf{F}}$ имеет вид

$$\begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix},$$

где c – корреляция линейной аппроксимации (u, v). Чему равны главные корреляции (U, V)?

Упражнение 11.8

Пусть (U, V) – множественное линейное свойство функции $\mathsf{F} : G \to G$, содержащее тривиальную линейную аппроксимацию. Квадрат нормы Фробениуса $\|\cdot\|_{\text{fr}}^2$ матрицы равен сумме квадратов ее сингулярных значений.

1. Докажите, что величина $\|\langle V, U\rangle_{\mathsf{F}}\|_{\text{fr}}^2 - 1$ равна емкости множественной линейной аппроксимации.
2. Воспользовавшись тем фактом, что норма Фробениуса не изменяется при унитарном преобразовании базиса, покажите, что $\|\langle V, U\rangle_{\mathsf{F}}\|_{\text{fr}}^2 - 1$ равна квадратичному евклидову расхождению.

Упражнение 11.9

Коалгеброй над $\mathbb{C}$ называется векторное пространство V с операцией *копроизведения* $\Delta: V \to V \otimes V$, которая удовлетворяет нескольким аксиомам. Например, $\mathbb{C}[G]$ является коалгеброй с копроизведением

$$\Delta(\delta_x) = \delta_{(x,x)}. \tag{11.1}$$

Следующие далее вопросы относятся к отображению $\Delta: \mathbb{C}[G] \to \mathbb{C}[G^2]$, определенному формулой (11.1). Обозначим id тождественную функцию на $\mathbb{C}[G]$.

1. Докажите, что Δ *коассоциативно*: $(\mathsf{id} \otimes \Delta) \circ \Delta = (\Delta \otimes \mathsf{id}) \circ \Delta$.
2. Покажите, что существует *коединица* $\epsilon : \mathbb{C}[G] \to \mathbb{C}$, для которой $(\mathsf{id} \otimes \epsilon) \circ \Delta = \mathsf{id}$ и $(\epsilon \otimes \mathsf{id}) \circ \Delta = \mathsf{id}$.
3. Морфизмом коалгебр называется линейное отображение $T : \mathbb{C}[G] \to \mathbb{C}[H]$, удовлетворяющее условию $\Delta_H \circ T = (T \otimes T) \circ \Delta_G$, где Δ_G – копроизведение на $\mathbb{C}[G]$, а Δ_H – копроизведение на $\mathbb{C}[H]$. Докажите, что для любого морфизма коалгебр $T : \mathbb{C}[G] \to \mathbb{C}[H]$ существует функция $\mathsf{F}: G \to H$ такая, что $T = T^{\mathsf{F}}$.

Структура коалгебры на $\mathbb{C}[G]$ играет важную роль в геометрическом подходе к криптоанализу вообще.

Приложение А. Нормальное распределение

В этом приложении собраны некоторые важные факты, касающиеся нормального распределения. Эти результаты используются в книге, особенно в главах 4, 6 и 7.

А.1. Одномерное нормальное распределение

Нормальные распределения – это семейство непрерывных распределений вероятностей. *Стандартное нормальное распределение* имеет функцию плотности

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

График этой функции показан на рис. А.1.

Функция распределения стандартного нормального распределения обозначается буквой Ф. По определению, она имеет вид

$$\Phi(x) = \int_{-\infty}^{x} \varphi(z)\, dz.$$

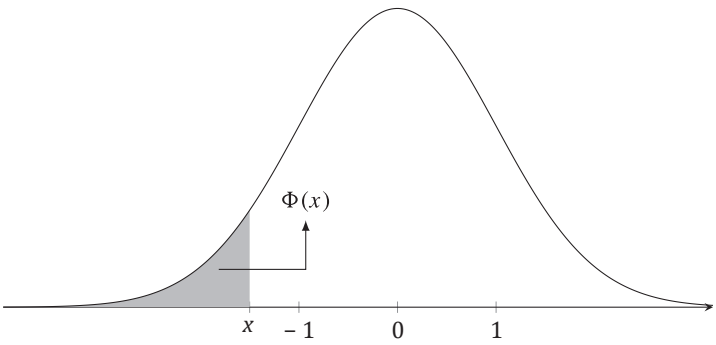


Рис. А.1. Функция плотности вероятности стандартного нормального распределения

Из симметрии ϕ следует, что среднее стандартного нормального распределения равно нулю. С помощью интегрирования по частям можно показать, что дисперсия равна 1.

Другие нормальные распределения получаются из стандартного путем масштабирования и сдвига. Если $\mathbf{x}$ – случайная величина со стандартным нормальным распределением, то функция распределения $\sigma\mathbf{x} + \mu$ переводит x в

$$\Phi\Big(\frac{x-\mu}{\sigma}\Big).$$

Эквивалентно можно сказать, что функция плотности $\sigma\mathbf{x} + \mu$ переводит x в

$$\frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

Среднее и дисперсия однозначно определяют каждый член семейства нормальных распределений. Поэтому разумно обозначить нормальное распределение со средним μ и дисперсией σ^2 $\mathcal{N}(\mu, \sigma^2)$.

Особая важность нормального распределения связана со следующей предельной теоремой.

Теорема А.1 (центральная предельная теорема). *Пусть* $\mathbf{x}_1, \mathbf{x}_2, \ldots$ – *последовательность независимых случайных величин на* $\mathbb{R}$ *со средним* μ *и дисперсией* σ^2. *В пределе при* $n \to \infty$ *распределение* $\sum_{i=1}^{n} \mathbf{x}_i/\sqrt{n}$ *сходится к* $\mathcal{N}(\mu, \sigma^2)$.

Одно из следствий теоремы А.1 – нормальная аппроксимация биномиального распределения. Точнее, биномиальное распределение с n испытаниями и вероятностью успеха p при больших n хорошо аппроксимируется распределением $\mathcal{N}(np, np(1-p))$.

А.2. Многомерное нормальное распределение

Стандартным многомерным нормальным распределением называется распределение вероятностей вектора, состоящего из d независимых стандартных нормальных распределений. Следовательно, его плотность равна

$$\varphi(x) = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}x_i^2} = \frac{1}{\sqrt{(2\pi)^d}}\, e^{-\frac{1}{2}\|x\|_2^2},$$

где $(x_1, \ldots, x_d)$ – элементы вектора $x \in \mathbb{R}^d$.

Другие многомерные нормальные распределения получаются из стандартного с помощью аффинных преобразований. А именно если $\mathbf{x}$ – случайная величина со стандартным многомерным нормальным распределением, то функция плотности вероятности $A\mathbf{x} + \mu$ переводит x в

$$\frac{1}{|\det A|}\frac{1}{\sqrt{(2\pi)^d}}\, e^{-\frac{1}{2}\|A^{-1}(x-\mu)\|_2^2},$$

где μ – вектор в $\mathbb{R}^d$, а A – обратимая матрица размера $d \times d$. Множитель $|\det A|$ – это якобиан преобразования $x \mapsto Ax + \mu$. Вектор μ является средним распределения $A\mathbf{x} + \mu$. Ковариационная матрица $A\mathbf{x} + \mu$ равна

$$\Sigma = \mathop{\mathbb{E}}_{\mathbf{x}} (A\mathbf{x})(A\mathbf{x})^{\mathsf{T}} = A\Big(\mathop{\mathbb{E}}_{\mathbf{x}} \mathbf{x}^{\mathsf{T}}\mathbf{x}\Big)A^{\mathsf{T}} = AA^{\mathsf{T}}.$$

Среднее μ и ковариационная матрица Σ однозначно определяют каждый член семейства многомерных нормальных распределений. Многомерное нормальное распределение со средним μ и ковариационной матрицей Σ обозначается $\mathcal{N}(\mu, \Sigma)$. Его функция плотности распределения переводит $x \in \mathbb{R}^d$ в

$$\frac{1}{\sqrt{|\det \Sigma|}} \frac{1}{\sqrt{(2\pi)^d}} e^{-\frac{1}{2}(x-\mu)^\intercal \Sigma^{-1}(x-\mu)}.$$

Если $\mathbf{x}$ – случайная величина с распределением $\mathcal{N}(\mu, \Sigma)$, то для любого вектора $v \in \mathbb{R}^d$ распределение $v^\intercal \mathbf{x}$ является одномерным нормальным распределением $\mathcal{N}(v^\intercal \mu, v^\intercal \Sigma v)$. Среднее получается из линейности математического ожидания, а дисперсия вычисляется по формуле

$$\mathop{\mathbb{V}}_{\mathbf{x}} v^\intercal \mathbf{x} = \mathop{\mathbb{E}}_{\mathbf{x}} \left(v^\intercal (\mathbf{x} - \mu)\right)^2 = v^\intercal \left(\mathop{\mathbb{E}}_{\mathbf{x}} (\mathbf{x} - \mu)(\mathbf{x} - \mu)^\intercal\right) v = v^\intercal \Sigma\, v.$$

Существует вариант центральной предельной теоремы для многомерных распределений.

Теорема А.2 (центральная предельная теорема). *Пусть* $\mathbf{x}_1, \mathbf{x}_2, \ldots$ – *последовательность независимых случайных величин на* $\mathbb{R}^d$ *со средним* μ *и ковариационной матрицей* Σ. *В пределе при* $n \to \infty$ *распределение* $\sum_{i=1}^n \mathbf{x}_i / \sqrt{n}$ *сходится к* $\mathcal{N}(\mu, \Sigma)$.

Приложение B. Краткий справочник по статистике

Формулы в табл. B.1 опираются на несколько аппроксимаций, которые перечислены в табл. B.2. Приведенные ниже замечания относятся к табл. B.1.

Одиночные аппроксимации. Для выборки без возвращения поделить аргумент Φ на $\sqrt{1 - q/2^n}$ (см. раздел 4.3, стр. 66).

Множественные аппроксимации. Если корреляции известны с точностью до знака, то заменить Cap(Λ) на (см. раздел 6.1.2, стр. 83)

$$\sqrt{\sum_{i=1}^{|\Lambda|} c_i^4}.$$

Многомерные аппроксимации. Если корреляции неизвестны и $\Lambda = \Lambda_{in} \oplus \Lambda_{out}$, то заменить $|\Lambda|$ на $|\Lambda_{out}|$, если доступны выбранные открытые тексты (см. раздел 6.2.3, стр. 88). Для многомерных линейных аппроксимаций с нулевой корреляцией заменить $|\Lambda|$ на $|\Lambda_{in}|^2 |\Lambda_{out}|$ (см. раздел 8.4, стр. 117).

Если корреляции или емкость зависят от ключа, то использовать следующую формулу:

$$\sum_{k \in \mathcal{K}} f_k \, P_S(k),$$

где f_k – частота k-го класса ключей. Эти частоты выводятся из априорного распределения ключа. Может потребоваться дополнительный анализ развертки ключа. Приведенную выше формулу можно модифицировать, так чтобы она учитывала ошибки модели (см. обсуждение в разделе 7.3.2 на стр. 104).

Таблица В.1. Основные статистические формулы для вероятности успеха P_S

	Корреляции	
	Известны	Неизвестны
Одиночная аппроксимация	$\Phi\left(\lvert c\rvert\sqrt{q} + \Phi^{-1}(P_F)\right)$	$\Phi\left(\lvert c\rvert\sqrt{q} + \Phi^{-1}(P_F/2)\right)$
Множественная аппроксимация	$\Phi\left(\sqrt{\mathrm{Cap}(\Lambda)\, q} + \Phi^{-1}(P_F)\right)$	$\Phi\left(\frac{\mathrm{Cap}(\Lambda)}{\sqrt{2\lvert\Lambda\rvert}}\, q + \Phi^{-1}(P_F)\right)$
Множественная с нулевой корреляцией	$\Phi\left(\frac{\sqrt{\lvert\Lambda\rvert}}{2^{n+\frac{1}{2}}}\, q + \Phi^{-1}(P_F)\right)$	/

Таблица В.2. Аппроксимации, используемые для формул из табл. В.1

	Корреляции	
	Известны	Неизвестны
Одиночная аппроксимация	нулевая корреляция для неправильных ключей q велико (нормальная аппроксимация) $c^2 \ll 1$ (постоянная дисперсия)	
	c фиксировано	c^2 фиксировано P_F мало
Множественная аппроксимация	все корреляции равны нулю для неверных ключей $q/\sqrt{\lvert\Lambda\rvert}$ велико (нормальная аппроксимация) $\lvert C^F_{v_i+v_j,\,u_i+u_j}\rvert \ll 1/\sqrt{\lvert\Lambda\rvert}$ когда $(u_i + u_j, v_i + v_j) \notin \Lambda$ $c_i c_j \ll 1$ (ковариации пренебрежимо малы)	
	$c_1, c_2 \ldots$ фиксированы	Cap(Λ) фиксирована qc_i^2 мало

Приложение C. Список блочных шифров

Таблица C.1. Список блочных шифров, упоминаемых в этой книге

Блочный шифр	Глава	Ссылка
3-Way	Глава 1, стр. 15	Joan Daemen, René Govaerts, and Joos Vandewalle (Dec. 1994). «A New Approach to Block Cipher Design». In: *FSE'93*. Ed. by Ross J. Anderson. Vol. 809. LNCS. Springer, Berlin, Heidelberg, pp. 18–32. doi: 10.1007/3-540-58108-1_2
Simon	Глава 2, стр. 39	Ray Beaulieu et al. (2013). *The SIMON and SPECK Families of Lightweight Block Ciphers*. Cryptology ePrint Archive, Report 2013/404. URL: *https://eprint.iacr.org/2013/404*
Rijndael	Глава 3, стр. 46	Joan Daemen and Vincent Rijmen (2020). *The Design of Rijndael – The Advanced Encryption Standard (AES)*. 2nd ed. Information Security and Cryptography. Springer, Berlin, Heidelberg. isbn: 978-3-662-60768-8. doi: 10.1007/978-3-662-60769-5
Speck	Глава 3, стр. 51	Ray Beaulieu et al. (2013). *The SIMON and SPECK Families of Lightweight Block Ciphers*. Cryptology ePrint Archive, Report 2013/404

Литература

Ashur, Tomer, Tim Beyne, and Vincent Rijmen (Apr. 2020). «Revisiting the Wrong-Key-Randomization Hypothesis». In: *Journal of Cryptology* 33.2, pp. 567–594. doi: 10.1007/s00145-020-09343-2.

Baignères, Thomas, Pascal Junod, and Serge Vaudenay (Dec. 2004). «How Far Can We Go Beyond Linear Cryptanalysis?» In: *ASIACRYPT 2004*. Ed. by Pil Joong Lee. Vol. 3329. LNCS. Springer, Berlin, Heidelberg, pp. 432–450. doi: 10.1007/978-3-540-30539-2_31.

Baignères, Thomas, Jacques Stern, and Serge Vaudenay (Aug. 2007). «Linear Cryptanalysis of Non Binary Ciphers». In: *SAC 2007*. Ed. by Carlisle M. Adams, Ali Miri, and Michael J. Wiener. Vol. 4876. LNCS. Springer, Berlin, Heidelberg, pp. 184–211. doi: 10.1007/978-3-540-77360-3_13.

Banik, Subhadeep et al. (Nov. 2015). «Midori: A Block Cipher for Low Energy». In: *ASIACRYPT 2015, Part II*. Ed. by Tetsu Iwata and Jung Hee Cheon. Vol. 9453. LNCS. Springer, Berlin, Heidelberg, pp. 411–436. doi: 10.1007/978-3-662-48800-3_17.

Beaulieu, Ray et al. (2013). *The SIMON and SPECK Families of Lightweight Block Ciphers*. Cryptology ePrint Archive, Report 2013/404. url: https://eprint.iacr.org/2013/404.

Beierle, Christof, Anne Canteaut, and Gregor Leander (2018). «Nonlinear Approximations in Cryptanalysis Revisited». In: *IACR Transactions on Symmetric Cryptology* 2018.4, pp. 80–101. issn: 2519-173X. doi: 10.13154/tosc.v2018.i4.80-101.

Beyne, Tim (Dec. 2018). «Block Cipher Invariants as Eigenvectors of Correlation Matrices». In: *ASIACRYPT 2018, Part I*. Ed. by Thomas Peyrin and Steven Galbraith. Vol. 11272. LNCS. Springer, Cham, pp. 3–31. doi: 10.1007/978-3-030-03326-2_1.

Beyne, Tim (Dec. 2021). «A Geometric Approach to Linear Cryptanalysis». In: *ASIACRYPT 2021, Part I*. Ed. by Mehdi Tibouchi and Huaxiong Wang. Vol. 13090. LNCS. Springer, Cham, pp. 36–66. doi: 10.1007/978-3-030-92062-3_2.

Beyne, Tim (June 2023). «A Geometric Approach to Symmetric-Key Cryptanalysis». PhD thesis. KU Leuven.

Biryukov, Alex, Christophe De Cannière, and Michaël Quisquater (2004). «On Multiple Linear Approximations». In: *Advances in Cryptology – CRYPTO 2004, 24th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15–19, 2004, Proceedings*. Ed. by Matthew K. Franklin. Vol. 3152. LNCS. Springer, pp. 1–22. doi: 10.1007/978-3-540-28628-8_1.

Blondeau, Céline and Kaisa Nyberg (2017). «Joint Data and Key Distribution of Simple, Multiple, and Multidimensional Linear Cryptanalysis Test Statistic and Its Impact to Data Complexity». In: *Designs, Codes and Cryptography* 82, pp. 319–349.

Bogdanov, Andrey et al. (Dec. 2012). «Integral and Multidimensional Linear Distinguishers with Correlation Zero». In: *ASIACRYPT 2012*. Ed. by Xiaoyun Wang

and Kazue Sako. Vol. 7658. LNCS. Springer, Berlin, Heidelberg, pp. 244–261. doi: 10.1007/978-3-642-34961-4_16.

Bogdanov, Andrey and Vincent Rijmen (2014). «Linear Hulls with Correlation Zero and Linear Cryptanalysis of Block Ciphers». In: *DCC* 70.3, pp. 369–383. doi: 10.1007/s10623-012-9697-z.

Bogdanov, Andrey and Elmar Tischhauser (Mar. 2014). «On the Wrong Key Randomisation and Key Equivalence Hypotheses in Matsui's Algorithm 2». In: *FSE 2013*. Ed. by Shiho Moriai. Vol. 8424. LNCS. Springer, Berlin, Heidelberg, pp. 19–38. doi: 10.1007/978-3-662-43933-3_2.

Bogdanov, Andrey and Meiqin Wang (Mar. 2012). «Zero Correlation Linear Cryptanalysis with Reduced Data Complexity». In: *FSE 2012*. Ed. by Anne Canteaut. Vol. 7549. LNCS. Springer, Berlin, Heidelberg, pp. 29–48. doi: 10.1007/978-3-642-34047-5_3.

Collard, Baudoin and Francëois-Xavier Standaert (Apr. 2009). «A Statistical Saturation Attack against the Block Cipher PRESENT». In: *CT-RSA 2009*. Ed. by Marc Fischlin. Vol. 5473. LNCS. Springer, Berlin, Heidelberg, pp. 195–210. doi: 10.1007/978-3-642-00862-7_13.

Collard, Baudoin, Francois-Xavier Standaert, and Jean-Jacques Quisquater (2007). «Improving the Time Complexity of Matsui's Linear Cryptanalysis». In: *Information Security and Cryptology – ICISC 2007: 10th International Conference, Seoul, Korea, November 29–30, 2007. Proceedings 10*. Springer, Berlin, Heidelberg, pp. 77–88. doi: 10.1007/978-3-540-76788-6_7.

Daemen, Joan (Mar. 1995). «Cipher and Hash Function Design Strategies Based on Linear and Differential Cryptanalysis». PhD thesis. KU Leuven.

Daemen, Joan, Renr̈e Govaerts, and Joos Vandewalle (Dec. 1994). «A New Approach to Block Cipher Design». In: *FSE'93*. Ed. by Ross J. Anderson. Vol. 809. LNCS. Springer, Berlin, Heidelberg, pp. 18–32. doi: 10.1007/3-540-58108-1_2.

Daemen, Joan, Renr̈e Govaerts, and Joos Vandewalle (Dec. 1995). «Correlation Matrices». In: *FSE'94*. Ed. by Bart Preneel. Vol. 1008. LNCS. Springer, Berlin, Heidelberg, pp. 275–285. doi: 10.1007/3-540-60590-8_21.

Daemen, Joan, Lars R. Knudsen, and Vincent Rijmen (Jan. 1997). «The Block Cipher Square». In: *FSE'97*. Ed. by Eli Biham. Vol. 1267. LNCS. Springer, Berlin, Heidelberg, pp. 149–165. doi: 10.1007/BFb0052343.

Daemen, Joan and Vincent Rijmen (Dec. 2001). «The Wide Trail Design Strategy». In: *8th IMA International Conference on Cryptography and Coding*. Ed. By Bahram Honary. Vol. 2260. LNCS. Springer, Berlin, Heidelberg, pp. 222–238. doi: 10.1007/3-540-45325-3_20.

Daemen, Joan and Vincent Rijmen (2020). *The Design of Rijndael – The Advanced Encryption Standard (AES)*. 2nd ed. Information Security and Cryptography. Springer, Berlin, Heidelberg. isbn: 978-3-662-60768-8. doi: 10.1007/978-3-662-60769-5.

Halmos, Paul R. (1958). *Finite-dimensional Vector Spaces*. 1st ed. Undergraduate Texts in Mathematics. Springer New York, NY.

Harpes, Carlo, Gerhard G. Kramer, and James L. Massey (May 1995). «A Generalization of Linear Cryptanalysis and the Applicability of Matsui's Piling-Up Lemma».

In: *EUROCRYPT'95*. Ed. by Louis C. Guillou and Jean-Jacques Quisquater. Vol. 921. LNCS. Springer, Berlin, Heidelberg, pp. 24–38. doi: 10.1007/3-540-49264-X_3.

Harpes, Carlo and James L. Massey (Jan. 1997). «Partitioning Cryptanalysis». In: *FSE'97*. Ed. by Eli Biham. Vol. 1267. LNCS. Springer, Berlin, Heidelberg, pp. 13–27. doi: 10.1007/BFb0052331.

Hermelin, Miia, Joo Yeon Cho, and Kaisa Nyberg (2008). «Multidimensional Linear Cryptanalysis of Reduced Round Serpent». In: *Information Security and Privacy, 13th Australasian Conference, ACISP 2008, Wollongong, Australia, July 7–9, 2008, Proceedings*. Ed. by Yi Mu, Willy Susilo, and Jennifer Seberry. Vol. 5107. LNCS. Springer, pp. 203–215. doi: 10.1007/978-3-540-70500-0_15.

Kaliski Jr., Burton S. and Matthew J. B. Robshaw (Aug. 1994). «Linear Cryptanalysis Using Multiple Approximations». In: *CRYPTO'94*. Ed. by Yvo Desmedt. Vol. 839. LNCS. Springer, Berlin, Heidelberg, pp. 26–39. doi: 10.1007/3-540-48658-5_4.

Knudsen, Lars R. and Matthew J. B. Robshaw (May 1996). «Non-Linear Approximations in Linear Cryptanalysis». In: *EUROCRYPT'96*. Ed. by Ueli M. Maurer. Vol. 1070. LNCS. Springer, Berlin, Heidelberg, pp. 224–236. doi: 10.1007/3-540-68339-9_20.

Kullback, Solomon and Richard A. Leibler (1951). «On Information and Sufficiency». In: *The Annals of Mathematical Statistics* 22.1, pp. 79–86.

Leander, Gregor et al. (Aug. 2011). «A Cryptanalysis of PRINTcipher: The Invariant Subspace Attack». In: *CRYPTO 2011*. Ed. by Phillip Rogaway. Vol. 6841. LNCS. Springer, Berlin, Heidelberg, pp. 206–221. doi: 10.1007/978-3-642-22792-9_12.

Leander, Gregor, Brice Minaud, and Sondre Rшnjom (Apr. 2015). «A Generic Approach to Invariant Subspace Attacks: Cryptanalysis of Robin, iSCREAM and Zorro». In: *EUROCRYPT 2015, Part I*. Ed. by Elisabeth Oswald and Marc Fischlin. Vol. 9056. LNCS. Springer, Berlin, Heidelberg, pp. 254–283. doi: 10.1007/978-3-662-46800-5_11.

Matsui, Mitsuru (May 1994a). «Linear Cryptanalysis Method for DES Cipher». In: *EUROCRYPT'93*. Ed. by Tor Helleseth. Vol. 765. LNCS. Springer, Berlin, Heidelberg, pp. 386–397. doi: 10.1007/3-540-48285-7_33.

Matsui, Mitsuru (Aug. 1994b). «The First Experimental Cryptanalysis of the Data Encryption Standard». In: *CRYPTO'94*. Ed. by Yvo Desmedt. Vol. 839. LNCS. Springer, Berlin, Heidelberg, pp. 1–11. doi: 10.1007/3-540-48658-5_1.

Nyberg, Kaisa (May 1995). «Linear Approximation of Block Ciphers (Rump Session)». In: *EUROCRYPT'94*. Ed. by Alfredo De Santis. Vol. 950. LNCS. Springer, Berlin, Heidelberg, pp. 439–444. doi: 10.1007/BFb0053460.

Schulte-Geers, Ernst (2013). «On CCZ-equivalence of Addition mod 2 *n*». In: *Designs, Codes and Cryptography* 66, pp. 111–127.

Selcëuk, Ali Aydin (Jan. 2008). «On Probability of Success in Linear and Differential Cryptanalysis». In: *Journal of Cryptology* 21.1, pp. 131–147. doi: 10.1007/s00145-007-9013-7.

Tardy-Corfdir, Anne and Henri Gilbert (Aug. 1992). «A Known Plaintext Attack of FEAL-4 and FEAL-6». In: *CRYPTO'91*. Ed. by Joan Feigenbaum. Vol. 576. LNCS. Springer, Berlin, Heidelberg, pp. 172–181. doi: 10.1007/3-540-46766-1_12.

Terras, Audrey (1999). *Fourier Analysis on Finite Groups and Applications*. London Mathematical Society Student Texts. Cambridge University Press, Cambridge.

Todo, Yosuke, Gregor Leander, and Yu Sasaki (Dec. 2016). «Nonlinear Invariant Attack – Practical Attack on Full SCREAM, iSCREAM, and Midori64». In: *ASIACRYPT 2016, Part II*. Ed. by Jung Hee Cheon and Tsuyoshi Takagi. Vol. 10032. LNCS. Springer, Berlin, Heidelberg, pp. 3–33. doi: 10.1007/978-3-662-53890-6_1.

Vaudenay, Serge (1996a). «An Experiment on DES Statistical Cryptanalysis». In: *CCS '96, Proceedings of the 3rd ACM Conference on Computer and Communications Security, New Delhi, India, March 14–16, 1996*. Ed. by Li Gong and Jacques Stearn. ACM, New York, pp. 139–147. doi: 10.1145/238168.238206.

Vaudenay, Serge (Mar. 1996b). «An Experiment on DES Statistical Cryptanalysis». In: *ACM CCS 96*. Ed. by Li Gong and Jacques Stern. ACM Press, New York, pp. 139–147. doi: 10.1145/238168.238206.

Wall'en, Johan (Feb. 2003). «Linear Approximations of Addition Modulo 2 *n*.». In: *FSE 2003*. Ed. by Thomas Johansson. Vol. 2887. LNCS. Springer, Berlin, Heidelberg, pp. 261–273. doi: 10.1007/978-3-540-39887-5_20.

Предметный указатель

Р

С

Т

У

Книги издательства «ДМК Пресс» можно купить оптом и в розницу
на складе издательства по адресу:
Москва, ул. Электродная, д. 2, стр. 12, офис 7,
тел. +7 (499) 322–19–38,
а также заказать на сайте **www.dmkpress.com**
с доставкой в любой регион РФ

Тим Бейн, Винсент Рэймен

Линейный криптоанализ

Главный редактор *Яценков В. С.*
editor@dmkpress.com

Перевод *Слинкин А. А.*
Корректор *Синяева Г. И.*
Верстка *Луценко С. В.*
Дизайн обложки *Трофимова С. В.*

Формат 70×100 1/16.
Гарнитура «PT Serif». Печать цифровая.
Усл. печ. л. 14,63. Тираж 200 экз.

Веб-сайт издательства: *www.dmkpress.com*

Криптография находит множество применений в повседневной жизни. Это руководство посвящено анализу безопасности (криптоанализу) фундаментальных блоков, на которых основаны криптографические приложения. Линейный криптоанализ рассматривается с математической точки зрения и сопровождается обзором наиболее влиятельных публикаций. Главы дополнены большим количеством примеров и упражнений, опирающихся на теорию и практику.

Книга охватывает следующие темы:

- **линейные приближения и следы;**
- **корреляционные матрицы;**
- **автоматический поиск;**
- **методы восстановления ключей;**
- **многомерный линейный криптоанализ;**
- **приближения нулевой корреляции и геометрический подход.**

Предварительные знания теории криптографии не требуются. Издание будет полезно как начинающим читателям, изучающим криптографию, так и опытным экспертам, применяющим ее на практике.

Тим Бейн – научный сотрудник в Лёвенском университете, Бельгия. Его исследования были отмечены наградами на различных конференциях по криптографии, а также научной премией Nokia Bell 2024 г.

Винсент Рэймен – профессор Лёвенского университета и адъюнкт-профессор Университета Бергена, Норвегия. Он также является соавтором расширенного стандарта шифрования (AES). Получил докторскую степень за разработку и криптоанализ блочных шифров.

CAMBRIDGE UNIVERSITY PRESS

ДМК ПРЕСС издательство

www.дмк.рф

ISBN 978-5-93700-474-1

9 785937 004741 >