Оглавление

Предисловие от издательства 9

Предисловие 10

Глава 1. Введение 13

1.1. Криптографические примитивы 13
1.1.1. Анализ 13
1.1.2. Проектирование 14
1.2. Линейные аппроксимации 15
1.2.1. Смещение 16
1.2.2. Таблицы линейной аппроксимации 17
1.3. Линейные следы и лемма о набегании знаков 19
1.4. Восстановление ключа 21
1.4.1. Алгоритм Мацуи 1 21
1.4.2. Алгоритм Мацуи 2 23
1.5. Оставшиеся проблемы 24
1.6. Историческая справка 24
1.7. Литература 25
1.8. Упражнения 25

Глава 2. Корреляционные матрицы 27

2.1. Корреляция случайной величины на $\mathbb{F}_2$ 27
2.2. Корреляция между булевыми функциями 28
2.3. Корреляционные матрицы 29
2.4. Корреляционные матрицы структурных функций 31
2.5. Линейные следы 33
2.6. Историческая справка 35
2.7. Литература 36
2.8. Упражнения 37

Глава 3. Оптимизация линейных следов 42

3.1. Метод ветвей и границ 42
3.1.1. Поиск в глубину 42
3.1.2. Метод Мацуи 43
3.2. Смешанно-целочисленное линейное программирование 46
3.2.1. Пример: шифр типа Rijndael 46
3.2.2. Построение модели 48
3.2.3. Решение модели 50
3.3. Выполнимость и выполнимость в теориях 50
3.3.1. Пример: шифр add-rotate-xor 51

3.3.2. Построение модели ... 52
3.3.3. Решение модели ... 53
3.4. Историческая справка ... 54
3.5. Литература ... 54
3.6. Упражнения ... 54

Глава 4. Статистика линейного криптоанализа ... 58
4.1. Статистический вывод ... 58
4.1.1. Статистические оценки ... 58
4.1.2. Проверка гипотез ... 59
4.2. Восстановление ключа с помощью проверки статистических гипотез ... 62
4.2.1. Известная корреляция ... 62
4.2.2. Неизвестная корреляция ... 64
4.3. Стратегии выборки ... 66
4.4. Восстановление ключа с использованием ранжирования ключей ... 67
4.5. Историческая справка ... 68
4.6. Литература ... 68
4.7. Упражнения ... 68

Глава 5. Методы восстановления ключа ... 69
5.1. Восстановление ключа по алгоритму 2 ... 69
5.2. Подход Мацуи ... 70
5.2.1. Однонаправленный случай ... 70
5.2.2. Двунаправленный случай ... 72
5.3. Метод быстрого преобразования Фурье ... 73
5.3.1. Циркулянтная структура ... 73
5.3.2. Умножение на циркулянтные матрицы ... 74
5.4. Историческая справка ... 76
5.5. Литература ... 76
5.6. Упражнения ... 77

Глава 6. Множественный линейный криптоанализ ... 78
6.1. Множественный линейный криптоанализ ... 78
6.1.1. Множественные линейные аппроксимации ... 78
6.1.2. Различители ... 81
6.2. Многомерный линейный криптоанализ ... 84
6.2.1. Многомерные линейные аппроксимации ... 84
6.2.2. Различители ... 86
6.2.3. Атаки с выбранным открытым текстом ... 87
6.3. Заключительные замечания ... 88
6.3.1. Восстановление ключа ... 88
6.3.2. Нахождение подходящих линейных аппроксимаций ... 89
6.4. Историческая справка ... 89
6.5. Литература ... 89
6.6. Упражнения ... 90

Глава 7. Оптимальная проверка статистических гипотез ... 94
7.1. Вероятностные меры ... 94

7.2. Простые гипотезы 95
7.2.1. Теория Неймана–Пирсона 96
7.2.2. Два многомерных нормальных распределения 97
7.2.3. Два распределения почти равны 98
7.3. Составные гипотезы 101
7.3.1. Коэффициенты Байеса 102
7.3.2. Гипотеза рандомизации с правильным ключом 102
7.3.3. Гипотеза рандомизации с неправильным ключом 104
7.4. Оптимальное восстановление ключа 106
7.5. Историческая справка 107
7.6. Литература 107
7.7. Упражнения 107

Глава 8. Аппроксимации с нулевой корреляцией 109

8.1. Идея 109
8.2. Нахождение аппроксимаций с нулевой корреляцией 110
8.3. Использование аппроксимаций с нулевой корреляцией 112
8.3.1. Одна аппроксимация 113
8.3.2. Несколько аппроксимаций 115
8.4. Статистический подход 116
8.5. Историческая справка 117
8.6. Литература 117
8.7. Упражнения 118

Глава 9. Различные обобщения 121

9.1. Точные свойства 121
9.1.1. Атаки с насыщением 121
9.1.2. Инвариантные подпространства 123
9.1.3. Нелинейные инварианты 125
9.2. Приближенные свойства 127
9.2.1. Статистическое насыщение 127
9.2.2. Нелинейные аппроксимации 128
9.2.3. Каркас проецирования 128
9.3. Историческая справка 129
9.4. Литература 129
9.5. Упражнения 130

Глава 10. Функции на абелевых группах 132

10.1. Линейная алгебра над полем $\mathbb{C}$ 132
10.1.1. Нормированные векторные пространства и двойственные им 133
10.1.2. Пространства со скалярным произведением 135
10.1.3. Сингулярное разложение 137
10.1.4. Тензорные произведения векторных пространств 137
10.2. Анализ Фурье на конечных абелевых группах 138
10.2.1. Характеры группы 139
10.2.2. Преобразование Фурье 141
10.2.3. Двойственность Понтрягина 143

10.3. Историческая справка 144
10.4. Литература 144
10.5. Упражнения 145

Глава 11. Геометрический подход 148
11.1. Геометрический взгляд 148
11.1.1. Криптоаналитические свойства 148
11.1.2. Распространение 149
11.1.3. Геометрия 151
11.2. Линейный криптоанализ 152
11.2.1. Корреляционные матрицы 152
11.2.2. Множественный линейный криптоанализ 154
11.3. Точное распространение 155
11.3.1. Прямое распространение 155
11.3.2. Обратное распространение 155
11.3.3. Нулевая корреляция 156
11.3.4. Инварианты 156
11.4. Приближенное распространение 157
11.4.1. Отображения аппроксимации 157
11.4.2. Геометрия 158
11.4.3. Принцип доминирующих следов 159
11.5. Историческая справка 160
11.6. Литература 160
11.7. Упражнения 160

Приложение A. Нормальное распределение 164

Приложение B. Краткий справочник по статистике 167

Приложение C. Список блочных шифров 169

Литература 170

Предметный указатель 174

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 $i = 1, ..., q$. 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 $(k_1, ..., k_r)$ 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 $\mathbb{F}_2^n$):

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

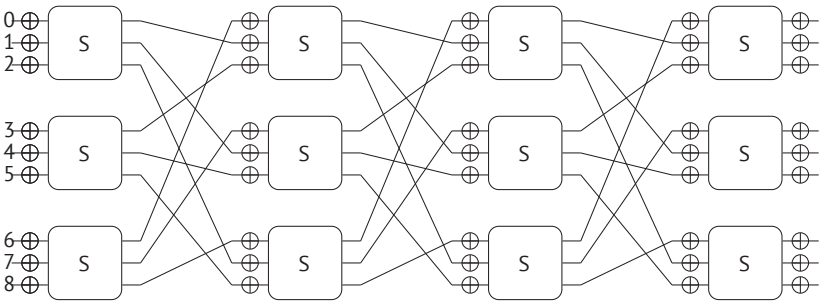


Рис. 1.1. Блочный шифр с размером блока 9 бит и четырьмя раундами

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

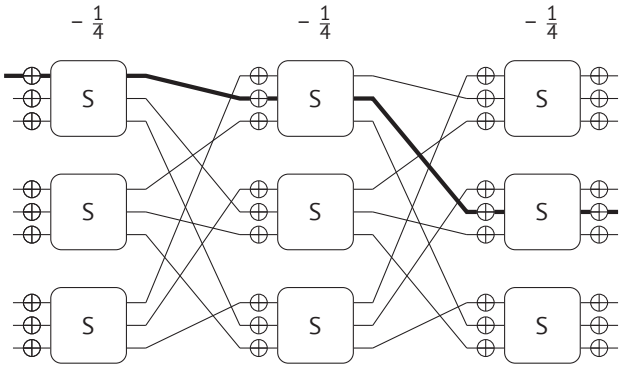


Рис. 1.2. Линейный след из примера 1.3

$\mathbf{a}_0 + \mathbf{b}_0 = 0$ 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

Конец ознакомительного фрагмента.
Приобрести книгу можно
в интернет-магазине
«Электронный универс»
e-Univers.ru